Optimal Transportation in the presence of a prescribed pressure field(Variational Problems and Related Topics)

Wolansky, G.

数理解析研究所講究録
1528 巻 2007 年 151-181

Optimal Transportation in the presence of a prescribed pressure field

G. Wolansky[1]

Department of mathematics, Technion, Haifa 32000, Israel

November 26, 2006

Abstract

The optimal (Monge-Kantorovich) transportation problem is discussed from several points of view. The Lagrangian formulation extends the action of the *Lagrangian* $L(v,x,t)$ from the set of orbits in $\mathbb{R}^n$ to a set of measure-valued orbits. The *Eulerian*, dual formulation leads an optimization problem on the set of sub-solutions of the corresponding Hamilton-Jacobi equation. Finally, the Monge problem and its Kantorovich relaxation are obtained by reducing the optimization problem to the set of measure preserving mappings and two point distribution measures subjected to an appropriately defined cost function.

In this paper we concentrate on mechanical Lagrangians $L=|v|^2/2+P(x,t)$ leading, in general, to a non-homogeneous cost function. The main results yield existence of a unique *flow* of homomorphisms which transport the optimal measure valued orbit of the extended Lagrangian, as well as the existence of an optimal solution to the dual Euler problem and its relation to the Monge- and Kantorovich formulations.

1 Introduction

1.1 Historical Background

The classical problem of optimal mass transportation was suggested by Monge in the 18'th century [M]: given a cost function $c(x,y)$ (originally, $c=|x-y|$) and a pair of Borel probability measures μ_0, μ_1 on (say) a common probability space Ω, minimize

$$\int c(x,\mathbf{T}(x))\mu_0(dx) \tag{M}$$

along all Borel mappings $\mathbf{T}:\Omega\to\Omega$ which transport μ_0 into μ_1 ($\mathbf{T}_{\#}\mu_0=\mu_1$), namely

$$\mu_0\left(\mathbf{T}^{-1}A\right)=\mu_1(A) \ \ \forall \text{ Borel sets } A\subset\Omega\ . \tag{1.1}$$

The Monge problem was revived in the last century. In particular, Kantorovich [K] introduced in 1942 a relaxation, reducing the Monge problem to a linear programming in a cone of two-point distributions over Ω whose marginals are μ_0,μ_1 respectively:

$$\min_{\lambda}\int\int c(x,y)\lambda(dx,dy) \quad ; \ \pi^{(0)}_{\#}\lambda=\mu_0\ ,\ \pi^{(1)}_{\#}\lambda=\mu_1 \tag{K}$$

[1]Acknowledgements: Research partially supported by Israel Science Foundation grant 406/05, founded by the Israeli Academy of Sciences and Humanities, and by the Fund for the Promotion of Research at the Technion.

Here $\pi^{(i)}, i = 0, 1$ are the natural projections of $\Omega \times \Omega$ on its factors. A particular attention is given to the *Wasserstein metrics*

$$W_p(\mu_0, \mu_1) = \left[\min_{\lambda} \int\int |x-y|^p \lambda(dx, dy) \quad ; \; \pi^{(0)}_{\#}\lambda = \mu_0 \; , \; \pi^{(1)}_{\#}\lambda = \mu_1\right]^{1/p} \tag{1.2}$$

where $p \geq 1$.

1.2 Objectives and main results

In general, if μ_0 contains an atom, then there is, in general, no deterministic mapping $\mathbf{T}$ of any type which maps μ_0 into μ_1, so there is no sense to compare the deterministic Monge problem (M) with the probabilistic Kantorovich problem (K). However, we may still consider the following alternative formulation in terms of *an optimal flow* with respect to some family of cost functions $c_{t_1,t_2} = J(x, y, t_1, t_2)$:

(F) : *Find a relaxed orbit* $\mu = \mu_{(t)}dt$ *and a flow of diffeomorphysms* $\mathbf{T}^{t_2}_{t_1} : \Omega \to \Omega$ *for* $t_1, t_2 \in (0, T)$ *such that*

(i) $\mathbf{T}^{t_2}_{t_1}$ *is the optimal Monge mapping with respect to* c_{t_1,t_2} *transporting* $\mu_{(t_1)}$ *to* $\mu_{(t_2)}$ *for any* $t_1, t_2 \in (0, T)$.

(ii) $\lim_{t\to 0} \mu_{(t)} = \mu_0$ *and* $\lim_{t\to T} \mu_{(t)} = \mu_1$ *in the weak sense of measures.*

(iii) The limits $\lim_{t\to T} \mathbf{T}^t_\tau =: \mathbf{T}^T_\tau$ *exists uniformly and* $\mathbf{T}^T_\tau$ *is a continuous mappings for any* $\tau \in (0, T)$.

It is feasible that, once a solution to the flow problem $\mathbf{F}$ is provided, a $c_{0,T}$ optimal solution to the Monge problem $\mathbf{M}$ with respect to μ_0, μ_1 exists by $\mathbf{T} = \lim_{\tau\to 0} \mathbf{T}^T_\tau$ provided the later limit exists as a Borel map.

Our starting point is the definition of a norm $||\mu||_p$ of a measure-valued orbit as the minimal $\mathbb{L}^p_\mu$-norm of the velocity fields $\boldsymbol{v}$ which satisfy the weak form of the continuity equation

$$\left\{\boldsymbol{v} = \boldsymbol{v}(x, t) \; ; \; \int_0^T \int_\Omega [\phi_t + \boldsymbol{v} \cdot \nabla_x \phi] \, \mu_{(t)}(dx)dt = 0 \quad ; \quad \forall \phi \in C^1_0(\Omega \times [0.T])\right\} \quad . \tag{1.3}$$

and

$$||\mu||_p := \left[\inf_{\boldsymbol{v}} \int_{\Omega\times[0,T]} |\boldsymbol{v}|^p \mu_{(t)}(dx)dt\right]^{1/p} \tag{1.4}$$

where the infimum is taken over all $\mu-$measurable vectorfield $\boldsymbol{v}$ satisfying (1.3). Denote the set for which $||\mu||_p < \infty$ as $\mathbf{H}_p$. This is a normed cone. In section 2 we shall indicate some of its properties and prove a compactness embedding of $\mathbf{H}_p$ (for $p > 1$) in a set of orbits which satisfies Holder continuity in an appropriate topology. In particular, the end conditions $\mu_0 := \mu_{(0)}$, $\mu_1 := \mu_{(T)}$ are uniquely defined for $\mu \in \mathbf{H}_p$ where $p > 1$.

In the rest of the paper we concentrate on the case $p = 2$. The connection between the cost function $c_{t_1,t_2} = J$ posted in formulation **(F)** above and the pressure P is as follows: The function $J = J_P$ is the action associated with the Lagrangian

$$J_P(x,y,t_1,t_2) = \inf_{\overline{x}} \left\{ \int_{t_1}^{t_2} \left[\frac{|\dot{\overline{x}}(t)|^2}{2} + P(\overline{x}(t),t) \right] dt \quad ; \; \overline{x} : [t_1,t_2] \to \Omega, \;\; \overline{x}(t_1) = x \; , \overline{x}(t_2) = y \right\} .$$

The main result of this paper, formulated in section 3, reveals a connection between the following approaches:

L: The Lagrangian approach: Minimize a Lagrangian L_P on the space of orbits $\mu \in \mathbf{H}_2$:

$$\mathcal{L} := \inf_{\mu} L_P(\mu) \; ; \; L_P(\mu) := \frac{1}{2}||\mu||_2^2 + \int_0^T \int_\Omega P\mu_{(t)}(dx)dt \; , \; \mu \in \mathbf{H}_2 \; , \; \mu_{(0)} = \mu_0, \mu_{(T)} = \mu_1 \, .$$

E: The Eulerian approach: Maximize on the set of velocity potentials ϕ

$$\mathcal{E} := \sup_{\phi} \left[\int_\Omega \phi(x,T)\mu_1(dx) - \int_\Omega \phi(x,0)\mu_0(dx) \right] \tag{1.5}$$

where the supremum is taken in the set of all functions $\phi = \phi(x,t)$ which are subsolutions of the Hamilton-Jacobi (HJ) equation (1.7) in a sense to be defined.

M: The Monge approach: Minimize on the set of mappings verifying (1.1)

$$\mathcal{M} := \inf_{\mathbf{T}} \left\{ \int_\Omega J_P(x,\mathbf{T}(x),0,T)\mu_0(dx) \;\; ; \;\; \mathbf{T}_{\#}\mu_0 = \mu_1 \right\}$$

K: The Kantorovich approach: Minimize on the set of 2-point probability measures with prescribed marginal

$$\mathcal{K} := \min_{\lambda} \left\{ \int \int J_P(x,y,0,T)\lambda(dx,dy) \quad ; \; \pi_{\#}^{(0)}\lambda = \mu_0 \; , \; \pi_{\#}^{(1)}\lambda = \mu_1 \right\}$$

Our first result reveals the relation between the above formulation: If $P \in C^1(\Omega \times [0,T])$ then

$$\mathcal{L} = \mathcal{E} = \mathcal{K}$$

holds for *arbitrary* (probability, Borel) end measures μ_0, μ_1. As discussed above, the Monge problem may not have a solution at all (e.g., if μ_0 contains an atomic measure and the set of transporting mappings $\mathbf{T}_{\#}\mu_0 = \mu_1$ is empty).

The second part of our main result shows the relation between the flow problem **(F)** and the Lagrangian formulation **L**. This is the relation between the optimal velocity field $\boldsymbol{v}$ realizing (1.4) and the induced flow

$$\frac{d}{dt}\mathbf{T}_{t_1}^t(x) = \boldsymbol{v}\left(\mathbf{T}_{t_1}^t, t\right) \; . \tag{1.6}$$

To elaborate, we shall prove

1) There exists a minimizer $\mu \in \mathbf{H}_2$ of $\mathbf{L}$ which satisfies the end conditions. This minimizer may be non-unique.

2) There exists a maximizer ψ of $\mathbf{E}$ which is a *Lipschitz* function on $\Omega \times [0,T]$ and satisfies the equation

$$\psi_t + \frac{1}{2}|\nabla_x \psi|^2 = P \tag{1.7}$$

almost everywhere. Again, such a maximizer may be non-unique.

3) The vector field $\boldsymbol{v} = \nabla_x \psi$ is defined *everywhere* on some relatively closed set $K_0 \subset \Omega \times (0,T)$ which contains the support of *any* minimal path μ of $\mathbf{L}$ given by (1).

Under some additional assumption on P (see Main Theorem in section 3) we also get

4) The vector field $\boldsymbol{v} = \nabla_x \psi$ is locally Lipschitz continuous on K_0.

5) The restriction of $\boldsymbol{v}$ to the support of *any* minimal orbit of $\mathbf{L}$ is uniquely determined.

6) The flow $\mathbf{T}$ induced by $\boldsymbol{v}$ (1.6) leaves K_0 invariant.

7) The flow $\mathbf{T}_{t_1}^{t_2}$ transports $\mu_{(t_1)}$ to $\mu_{(t_2)}$ for any minimizer μ of $\mathbf{L}$ and any $t_1, t_2 \in (0,T)$. Moreover, it is an optimal Monge transport with respect to the action $J_P(\cdot,\cdot,t_1,t_2)$.

8) The maps $\lim_{\tau\to T} \mathbf{T}_t^\tau := \mathbf{T}_t^T : \Omega \to \Omega$ and $\lim_{\tau\to 0} \mathbf{T}_\tau^t := \mathbf{T}_0^t : \Omega \to \Omega$ exist and are continuous for any $t \in (0,T)$. Moreover, $[\mathbf{T}_t^T]_\#$ (res. $[\mathbf{T}_0^t]_\#$) is an optimal Monge map with respect to the action $J_P(,,t,T)$ (res. $J_P(\cdot,\cdot,0,t)$) transporting $\mu_{(t)}$ to μ_1 (res. μ_0 to $\mu_{(t)}$).

9) If $\lim_{t\to T} \mathbf{T}_0^t := \overline{T}$ exists as a Borel map, then $\overline{\mathbf{T}}$ transports μ_0 to μ_1 and is an optimal solution of the Monge problem $\mathbf{M}$. In this case

$$\mathcal{M} = \mathcal{L} = \mathcal{E} = \mathcal{K}$$

A particular case is the *pressureless* flow $P \equiv 0$. Here the optimal potential satisfies

$$\psi_t + \frac{1}{2}|\nabla_x \psi|^2 = 0 \tag{1.8}$$

and the associated action is

$$J_0(x,y,t_1,t_2) = \frac{|x-y|^2}{2(t_2-t_1)} ,$$

reducing the Monge-Kantorovich problem to the Wasserstein metric W_2 for quadratic costs (1.2). The associated flow, claimed in (6), is given in this case by

$$\mathbf{T}_{t_1}^{t_2}(x) = x + (t_2 - t_1)\nabla_x \psi(x,t_1)$$

where $\nabla_x \psi$ is defined and Lipschitz *everywhere.* In particular it follows that, for a quadratic cost, an optimal Monge map $\mathbf{T}_\#\mu_0 = \mu_1$ exists and is unique provided $\nabla_x \psi(x,0)$ is μ_0

measurable.[2] In this case, Brenier representation $\mathbf{T} = \mathbf{T}_0^T = \nabla_x \Phi$ of the optimal map [B] is recovered via

$$\Phi(x) = x^2/2 + T\psi(x,0) \quad .$$

The connection between the Monge-Kantorovich problem in the quadratic case and the flow problem $\mathbf{L}$ ($P \equiv 0$), as well as the dual relation $\mathbf{E}$ together with the Hamilton-Jacobi equation (1.8) was indicated by several authors (see [BB], [BBG])[3] as well as in the excellent monograph of Villani [V]. However, to the best of my knowledge, the existence and uniqueness result for the flow $\mathbf{T}_{t_1}^{t_2}$ *without any regularity assumptions* on the end measures μ_0, μ_1 is new even in the case $P \equiv 0$. In fact, the existence and uniqueness of the flow holds even if there is no optimal Monge map.

In section 4 we shall start to develop the tools needed for the proof of our main results. Section 4.1 deals with a dual formulation for the norm $||\mu||_2$ for an orbit of measure $\mu = \mu_{(t)}dt \in \mathbf{H}_2$. It follows that

$$||\mu||_2 = \sqrt{\sup\left[\frac{(\int\int \phi_t \mu(dxdt))^2}{\int\int |\nabla_x \phi|^2 \mu(dxdt)}\right]}$$

where the supremum is taken on the set of test functions $\phi(x,t) = \phi \in C_0^1(\Omega \times [0,T])$. An equivalent definition turns out to be

$$\frac{1}{2}||\mu||_2^2 = \sup_{\phi,P}\left\{-\int_{\Omega_I} P(x,t)\mu(dxdt) - \int_\Omega \phi(x,0)\mu_0(dx) + \int_\Omega \phi(x,T)\mu_1(dx)\right\} \tag{1.9}$$

where the infimum above is on the pairs of "velocity potentials" $\phi \in C^1(\Omega \times [0,T])$ and "pressures" $P = P(x.t)$ which are related via the Bernulli-type (or Hamilton-Jacobi) equation (1.7). In case of a *prescribed* pressure P (as in this paper), this identity reveals the relation between the Lagrangian formulation $\mathbf{L}$ and the Eulerian one $\mathbf{E}$. In section 4.2 we imply a dual formulation to a strict convex perturbation of the Lagrangian $\mathbf{L}$, leading to an approximation of the Euler formulation $\mathbf{E}$, to be used in the proof of the main result.

For the proof of the main result we shall also need a series of auxiliary Lemmas and definitions related to the Hamilton-Jacobi equation. In subsection 5.1 we list these definitions and Lemmas, concerning forward (maximal), backward (minimal) and reversible solutions of the Hamilton-Jacobi equation, which are essential to the proof of the main results. The proofs of the Lemmas are given in [W1]. In 5.2 we utilize these results for the proofs of our main Theorem.

In the rest of the paper we shall restrict ourselves to the flat torus $\Omega = \mathbb{R}^n/\mathbb{Z}^n$. The reason is that we wish to avoid compactness problems originated from measures on $\mathbb{R}^n$, on the one hand, and the boundary conditions for the Hamilton-Jacobi equation required in case of a bounded domain $\Omega \subset \mathbb{R}^n$. The flat torus is the simplest example in the sense that it is compact manifold with no boundary, on the one hand, and it inherits the Euclidean geometry from $\mathbb{R}^n$ on the other. Any function (or probability measure) on Ω is understood

[2]Since $\psi(,0)$ is a Lipschitz function, $\nabla_x\psi(x,0)$ is a measurable function defined a.e, so we recover the existence of an optimal map if μ_0 is a continuous w.r to Lebesgue measure.

[3]I wish to thank Prof. D. Kinderlehrer for turning my attention to these publications.

as a periodic function (or periodic, normalized per-period measure) on $\mathbb{R}^n$, unless otherwise is *explicitly specified.* In particular, a mapping $\mathbf{T} : \Omega \to \Omega$ is understood as a mapping on the covering $\mathbb{R}^n$ which satisfies $\mathbf{T}(x+\boldsymbol{z}) = \mathbf{T}(x) + \boldsymbol{z}$ for any $x \in \mathbb{R}^n$ and any $\boldsymbol{z} \in \mathbb{Z}^n$.

List of symbols and definitions

- $\Omega := \mathbb{R}^n/\mathbb{Z}^n$.
- $I = [0,T]$; $I_0 = (0,T)$
- $\Omega_I = \Omega \times I$, $\Omega_{I_0} = \Omega \times I_0$.
- LIP_l is the set of all locally Lipschitz functions in $\Omega)I_0$.
- $\mathcal{M}$ is the set of all probability Borel measures supported in Ω.
- $\mathcal{M}_I$ is the set of all Borel probability measures supported on Ω_I which are decomposable as $\mu \in \mathcal{M}_I \Longleftrightarrow \mu = \mu_{(t)}dt$ where $\mu_{(t)} \in \mathcal{M}$ a.e. $t \in I$.
- if μ is Lebesgue continuous measure, then $\rho_\mu \in \mathbb{L}^1(\Omega_I)$ is the density of μ.
- $\pi^{(0)}$ (res. $\pi^{(1)}$) is the natural projection of $\Omega \times \Omega$ on its first (res. second) factor Ω.
- For any pair $\mu_0, \mu_1 \in \mathcal{M}$, the Wasserstein-p metric is defined by

  $$W_p(\mu_0, \mu_1) := \inf_\lambda \int_\Omega \int_\Omega |x-y|^p \lambda(dxdy)$$

  where the infimum is on all probability measures on $\Omega \times \Omega$ such that $\pi^{(0)}_\# \lambda = \mu_0$, $\pi^{(1)}_\# \lambda = \mu_1$.
- $\mathbb{E}_\mu(\psi) := \int_0^T \int_\Omega \psi(x,t)\mu_{(t)}(dx)dt$. Likewise, $\mathbb{E}_{\mu_{(t)}}(\psi) = \int_\Omega \psi(x,t)\mu_{(t)}(dx)$.
- A lifting ν of $\mu \in \mathcal{M}_I$ is a Borel measure on $\Omega_I \times \mathbb{R}^n$ such that

  $$\mathbb{E}_\nu(\psi) := \int_0^T \int_\Omega \int_{\mathbb{R}^n} \psi(x,t)\nu(dxdtdv) = \mathbb{E}_\mu(\psi) \;\; ; \; \mathbb{E}_\nu(\psi_t + v \cdot \nabla_x \psi) = 0$$

  for all $\psi \in C_0^1(\Omega_I)$.

2 A metric space for measure's orbits

We start with the following

Definition 2.1. *Let $\mu \in \mathcal{M}_I$. Then $\mu \in \mathbf{H}_p(I, \mathcal{M})$ if there exists a lifting ν of μ such that $\mathbb{E}_\nu(|v|^p) < \infty$. We shall also define the $\mathbf{H}_p$ norm of $\mu \in \mathbf{H}_p$ by:*

$$||\mu||_p = \inf_\nu \left[\mathbb{E}_\nu(|v|^p)\right]^{1/p}$$

where the infimum is taken over all liftings of μ.

Lemma 2.1. *$\mathbf{H}_p$ is complete and locally compact under the weak C^* topology if $p > 1$. That is, for any bounded sequence μ_n in $\mathbf{H}_p$ we can extract a subsequence which converges in $C^*(\Omega_I)$ to some $\mu \in \mathbf{H}_p$. In addition:*

$$\lim_{n\to\infty} ||\mu_n||_p \geq ||\mu||_p \ .$$

Proof. By definition there exists a set of liftings ν_n corresponding to μ_n. Moreover, this sequence can be chosen so that $\mathbb{E}_{\nu_n}(|v|^p) < C$, so ν_n and $v\nu_n$ are tight on $\Omega_I \times \mathbb{R}^n$ (since $p > 1$ and Ω_I is compact). Hence the weak limit ν of ν_n is a lifting of the weak limit μ of μ_n, and $\mathbb{E}_\nu(|v|^p) < C$, hence $\mu \in \mathbf{H}_p$. The same argument also yields the lower-semi-continuity of $\mathbf{H}_p$. □

Lemma 2.2. *If $\mu = \mu_{(t)}dt \in \mathbf{H}_p$, $p > 1$ then the map $t \to \mu_{(t)}$ is a Holder $(p-1)/p$ continuous function from I into $\mathcal{M}$ with respect to the weak (C^*) topology equipped with the Wasserstein-1 norm W_1:*

$$W_1(\mu_0, \mu_1) = \sup_{|\nabla\phi|\leq 1} \int_\Omega \phi(\mu_1(dx) - \mu_0(dx)) \ . \tag{2.1}$$

Proof. We know that an optimal lifting ν exists for $\mu \in \mathbf{H}_p$. The measure ν can be decomposed, by the Theorem of measure's decomposition [AFP], into $\nu = \mu_{(t)}(dx)\nu_{x,t}(dv)dt$, for μ a.a. (x,t). We may define now the velocity field

$$\boldsymbol{v}(x,t) = \mathbb{E}_{\nu_{x,t}}(v)$$

for μ a.a. (x,t). It follows that $\boldsymbol{v} \in \mathbb{L}^p_\mu$ and, moreover,

$$||\mu||_p = \left[\int_{\Omega_I} |\boldsymbol{v}|^p \mu(dxdt)\right]^{1/p} .$$

By assumption:

$$\int_I\int_\Omega \frac{\partial\phi}{\partial t}\mu_{(t)}(dx)dt = -\int_I\int_\Omega \boldsymbol{v}\cdot\nabla_x\phi\mu_{(t)}(dx)dt \tag{2.2}$$

where $\phi = \phi(x,t)$ is in $C_0^1(\Omega_I)$. Let $\phi(x,t) = h(t)\Phi(x)$ with $\Phi \in C^1(\Omega)$ and $h \in C_0^1(I)$. Then $f_\Phi(t) := \int_\Omega \Phi(x)\mu_{(t)}(dx)$ satisfies

$$\int_0^T f_\Phi(t)h'(t)dt = -\int_0^T h(t)\int_\Omega \nabla\Phi \cdot \boldsymbol{v}\mu_{(t)}(dx)dt \ .$$

By Holder inequality

$$\int_0^T h(t)\int_\Omega \nabla\Phi \cdot \boldsymbol{v}\mu_{(t)}(dx)dt \leq |\nabla\Phi|_\infty \left[\int_0^T h^q(t)dt\right]^{1/q} ||\mu||_p$$

with $q = p/(p-1)$. It follows that $f_\Phi \in \mathbb{W}^{1,p}(I)$ and, moreover, $||f_\Phi||_{1,p} \leq C||\Phi||_{1,\infty}$. This implies the result by Sobolev imbedding together with the dual formulation of the W_1 norm (2.1). □

Given μ_0 and $\mu_1 \in \mathcal{M}$, define the set

$$\Lambda_p(\mu_0,\mu_1) := \left\{\mu = \mu_{(t)}dt \in \mathbf{H}_p \ ; \ \mu_{(0)} = \mu_0 \ , \ \mu_{(T)} = \mu_1; \right\} \ .$$

Corollary 2.1. *The set $\Lambda_p(\mu_0,\mu_1)$ where $p > 1$ is closed and locally compact in $C(I;C^*(\Omega))$.*

Similar versions of the Lemma and Proposition below can be found in [Am]. We also note that Proposition 2.1 in the case $p = 2$ is a special case of our main Theorem (see section 3).

Lemma 2.3. (Regularization Lemma): *If $\mu \in \mathbf{H}_p$ then there exists a sequence $\mu^\varepsilon \in \mathbf{H}_p$ of smooth density so that $\mu = \lim_{\varepsilon\to0}\mu^\varepsilon$ holds in $C^*(\Omega_I)$ and, moreover,*

$$\lim_{\varepsilon\to0} ||\mu^\varepsilon||_p = ||\mu||_p \ .$$

In addition, for any $t_0, t_1 \in I$,

$$\lim_{\varepsilon\to0} W_p(\mu^\varepsilon_{t_0}, \mu^\varepsilon_{t_1}) = W_p(\mu_{(t_0)}, \mu_{(t_1)}) \ .$$

We next consider the relation between $\mathbf{H}_p$ and the optimal solution of the Kantorovich problem.

Proposition 2.1. *Assume $p \geq 1$. Let $\mu_0, \mu_1 \in \mathcal{M}$. Then $\Lambda_p(\mu_0,\mu_1) \neq \emptyset$. and*

$$\inf_{\mu\in\Lambda_p(\mu_0,\mu_1)} ||\mu||_p = W_p(\mu_0,\mu_1) \ .$$

The proof is similar to the proof of Theorem 4.2 of Ambrosio [Am] for the metric case $(p = 1)$.

We note that Corollary 2.1 is *not* valid in the case $p = 1$. To see it, consider the measure:

$$\mu = \sum_j \alpha_j(t)\delta_{(x-x_j(t))}dt$$

where $x_j = x_j(t) \in C^1(I;\Omega)$ and $\alpha_j \in C^1_+(I,\mathbb{R})$ such that $\sum_j \alpha_j(t) = 1 \ \forall t \in I$. We can approximate μ by a sequence of measures $\mu_m \in \Lambda_1(\mu_0,\mu_1)$ as follows: For each $m \in \mathbb{N}$ consider the division $t_k^{(m)} = k/m$, $0 \le k \le m$ of I. Let $\lambda_{m,k}$ be the optimal solution of Kantorovich problem due to $W_1(\mu_{(t_k^{(m)})}, \mu_{(t_{k+1}^{(m)})})$, and $\mathbf{T}_{m,k}^{(t)} := \mathbf{Id} + (t - t_k^{(m)})\,[\mathbf{T}_{m,k} - \mathbf{Id}]\,/(t_{k+1}^{(m)} - t_k^{(m)})$. Define μ_m as follows:

$$\mu_{m,(t_k)} = \mu_{(t_k)} \quad ; \quad \mu_{m,(t)} = \mathbf{T}_{m,k,\#}^{(t)}\mu_{m,(t_k)} \quad ; \quad t_k^{(m)} \le t \le t_{k+1}^{(m)} .$$

Then, by Proposition 2.1, μ_m are bounded in $\mathbf{H}_1$ and $\mu_m \to \mu$. However, $\mu \notin \mathbf{H}_1$ unless α_j are constants in t. To see it, note that the continuity equation takes the form

$$0 = \sum_j \int_I (\alpha_j(t)\phi_t(x_j(t),t) + v_j(t)\cdot\nabla_x\phi(x_j(t),t)dt = \sum_j \int_I -\dot{\alpha}_j\phi(x_j(t),t) + [v_j(t) - \dot{x}_j]\cdot\nabla_x\phi(x_j(t),t)dt$$

where $v_j(t)$ are the velocities attributed to x_j. It is evident that, unless $\dot{\alpha}_j \equiv 0$, for any possible choice of v_j one can find $\phi = \phi(x,t)$ for which the integral on the right does not vanish.

3 Main results

Let the pressure $P = P(x,t) \in C^1(\Omega_I)$ and the associated action:

$$L_P(\mu) := \frac{1}{2}||\mu||_2^2 + \int_{\Omega_I} P\mu(dxdt) \quad ; \quad \mu \in \mathbf{H}_2 . \tag{3.1}$$

Let us recall the definition of the action J_P:

$$J_P(x,y,t_1,t_2) = \inf_{\overline{x}} \left\{ \int_{t_1}^{t_2} \left[\frac{|\dot{\overline{x}}(t)|^2}{2} + P(\overline{x}(t),t) \right] dt \quad ; \ \overline{x} : [t_1,t_2] \to \Omega, \ \ \overline{x}(t_1) = x \ , \overline{x}(t_2) = y \right\} . \tag{3.2}$$

Remark: *Note that J_P is not a function on $\Omega = \mathbb{R}^n/\mathbb{Z}^n$ in each of the variables x, y, separately. However, for each $\mathbf{q} \in \mathbb{Z}^n$ and each $x, y \in \mathbb{R}^n$, $t_1, t_2 \in I$, $J_P(x+\mathbf{q}, y+\mathbf{q}, t_1, t_2) = J_P(x,y,t_1,t_2)$.*

Definition 3.1. *(**L**) (the relaxed Lagrangian):*

$$\mathcal{L}(\mu_0,\mu_1) := \inf_{\mu \in \Lambda_2(\mu_0,\mu_1)} L_P(\mu) .$$

Definition 3.2. *(**M**). (the Monge problem):*

$$\mathcal{M}(\mu_0,\mu_1) := \inf_{\mathbf{T}_\#\mu_0 = \mu_1} \int_\Omega J_P(x,\mathbf{T}(x),0,T)\mu_0(dx) .$$

Definition 3.3. *(**K**). (the Kantorovich problem):*

$$\mathcal{K}(\mu_0,\mu_1) := \inf_\lambda \int_\Omega J_P(x,y,0,T)\lambda(dxdy)$$

among all probability measures on $\Omega \times \Omega$ with the same Ω marginals μ_0, μ_1.

We now introduce the Hamilton-Jacobi (HJ) equation

$$\frac{\partial \phi}{\partial t} + \frac{1}{2}|\nabla_x \phi|^2 = P \ . \tag{3.3}$$

Let us denote the set of *classical sub-solutions* of the H-J equation as

$$\Lambda_P^* := \left\{ \phi \in C^1(\Omega_I) \ \ ; \ \ \phi_t + \frac{1}{2}|\nabla_x \phi|^2 \leq P \ \forall (x,t) \ \text{ in } \Omega_I \right\} \ .$$

For our purpose we need a generalization of the concept of a classical sub-solution. The concept of *viscosity sub-solution* (see, e.g. [E]) is too restrictive for us. So, we define a *generalized sub-solution* of the H-J equation as follows:

The set of generalized sub solution of the H-J equation is given by

$$\begin{aligned} \overline{\Lambda}_P^* := \{ \phi \in LIP(\Omega_I) \ ; \ \ \forall \overline{x} \in C^1(I;\Omega), \\ \frac{d}{dt}\phi(\overline{x}(t),t) \leq \frac{1}{2}\left|\dot{\overline{x}}(t)\right|^2 + P(\overline{x}(t),t) \ holds \ for \ Lebesgue \ \ a.e \ t \in I \right\} \end{aligned} \tag{3.4}$$

Remark (i): Note that $\phi(\overline{x}(t),t)$ is a Lipschitz function on I if ϕ is Lipschitz and $\overline{x} \in C^1(I)$. Hence it is a.e. differentiable (as a function of t) on I by Rademacher's Theorem (see, e.g., [E]).

Remark (ii): It is not difficult to see that any classical sub-solution is also generalized sub solution, so $\Lambda_P^* \subset \overline{\Lambda}_P^*$. The concept of generalized sub-solution is more general than that of a viscosity sub-solution. The relation between generalized sub-solutions and viscosity (and anti-viscosity) sub-solutions is discussed in section 5.1.

Definition 3.4. *(***E***): (The Euler formulation):*

$$\mathcal{E}(\mu_0,\mu_1) := \sup_{\phi \in \overline{\Lambda}^*(P)} \left\{ \int_\Omega \phi(x,T)\mu_1(dx) - \int_\Omega \phi(x,0)\mu_0(dx) \right\} \ .$$

We now state our main result:

Main Theorem:
Assume $P \in C^1(\Omega_I)$. For any μ_0, $\mu_1 \in \mathcal{M}$:

$$\mathcal{K}(\mu_0,\mu_1) = \mathcal{L}(\mu_0,\mu_1) = \mathcal{E}(\mu_0,\mu_1) \ . \tag{3.5}$$

There exists minimizers $\mu \in \Lambda_2(\mu_0,\mu_1)$ of **L** *(Definition 3.1) and a maximizer $\psi \in \overline{\Lambda}_P^*$ of* **E** *(Definition 3.4) such that*

$$\psi_t + \frac{1}{2}|\nabla_x \psi|^2 = P \quad ; \ a.e \ \ on \ \Omega_I \ . \tag{3.6}$$

Assume, in addition, there exists $C(t) > 0$ on I_0 so that $P(x,t) - C(t)|x|^2$ is a concave function on $\mathbb{R}^n$ for any $t \in I_0$. Then, for maximizer ψ of $\mathcal{E}(\mu_0,\mu_1)$, there exists a closed set $K \subset \Omega_I$ such that

i) *The restriction of* ψ *to* $K_0 := K \cap \Omega_{I_0}$ *is continuously differentiable, the equality* (3.6) *holds* for any $(x.t) \in K_0$ and $\nabla_x \psi$ is Locally Lipschitz continuous on K_0.

ii) Let v be a Lipschitz extension of $\nabla_x \psi$ to Ω_{I_0}. Let $\mathbf{T} = \mathbf{T}_{t_1}^{t_2}$ be the flow generated by v. Then K_0 is invariant under this flow.

iii) A minimizer $\mu \in \Lambda_2(\mu_0, \mu_1)$ of $\mathbf{L}$ is not necessarily unique. However, any such minimizer is supported in K and the vectorfield $v = \nabla_x \psi$ is uniquely defined on the support of any such minimizer.

iv) Any such minimizer is transported by the flow $\mathbf{T}$, that is

$$\left[\mathbf{T}_{t_1}^{t_2}\right]_{\#} \mu_{(t_1)} = \mu_{(t_2)}$$

holds for any $t_1, t_2 \in (0, T)$. Moreover, if $t_1 = 0$ (res. $t_2 = T$) then $\mathbf{T}_0^t := \lim_{\tau \to 0} \mathbf{T}_\tau^t$ (res. $\mathbf{T}_t^T = \lim_{\tau \to T} \mathbf{T}_t^\tau$) are continuous maps transporting μ_0 to $\mu_{(t)}$ (res. $\mu_{(t)}$ to μ_1).

v) The map $\mathbf{T}_{t_1}^{t_2}$ are optimal with respect to the cost function $c(x, y) = J_P(x, y, t_1, t_2)$ and the measures $\mu_{(t_1)}$, $\mu_{(t_2)}$, where either $t_1 \in I$, $t_2 \in I_0$ or $t_1 \in I_0$, $t_2 \in I$.

vi) If $P \equiv 0$ then the optimal solution ψ of $\mathbf{E}$ (Definition 3.4) is in $C_{loc}^{1,1}(\Omega_{I_0})$. In particular, the flow $\mathbf{T}$ can be defined *anywhere* in terms of ψ as

$$\mathbf{T}_{t_1}^{t_2}(x) = x + (t_2 - t_1)\nabla_x \psi(x, t_1) \ , \forall t_1 < t_2 \in I_0 \quad , \forall x \in \Omega \ .$$

4 Dual representation

The key duality argument for minimizing convex functionals under affine constraints is summarized in the following proposition whose proof is given in the appendix:

Proposition 4.1. *Let* $\mathbf{C}$ *a real Banach space and* $\mathbf{C}^*$ *the its dual. Denote the duality* $\mathbf{C} \div \mathbf{C}^*$ *relation by* $< c^*, c > \in \mathbb{R}$. *Let* $\mathbf{Z}$ *a subspace of* $\mathbf{C}$ *and* $h \in \mathbf{C}^*$. *Let* $\mathbf{Z}^* \subset \mathbf{C}^*$ *given by the condition* $z^* \in \mathbf{Z}^*$ *iff* $< z^* - h, z >= 0$ *for any* $z \in \mathbf{Z}$. *Let* $\mathcal{F} : \mathbf{C}^* \to \mathbb{R} \cup \{\infty\}$ *a convex function and*

$$I := \inf_{c^* \in \mathbf{Z}^*} \mathcal{F}(c^*) \ .$$

Assume further that $\overline{A}_0 := \{c^* \in \mathbf{C}^* \ ; \ \mathcal{F}(c^*) \le I\}$ *is compact (in the* $*-$ *topology of* $\mathbf{C}^*$*). Then*

$$\sup_{z \in \mathbf{Z}} \inf_{c^* \in \mathbf{C}^*} \left[\mathcal{F}(c^*) - < c^*, z > + < h, z >\right] = I \ .$$

In particular, both sides equal ∞ *if* $\mathbf{Z}^* = \emptyset$.

4.1 Dual representation of $\mathbf{H}_2$

We shall apply Proposition 4.1 were the space $\mathbf{C}$ is all the continuous functions $q = q(x,t,v)$ on $\Omega_I \times \mathbb{R}^n$ subjected to:

$$\|q\| := \sup_{(x,t,v)\in\Omega_I\times\mathbb{R}^n} \left\{ \frac{|q(x,t,v)|}{1+|v|^2} \right\} < \infty . \tag{4.1}$$

The dual space $\mathbf{C}^*$ contains all finite Borel measures ν on $\Omega_I \times \mathbb{R}^n$ of finite second moments:

$$\int_{\Omega_I\times\mathbb{R}^n} |\nu|(dxdtdv) < \infty \;\; ; \;\; \int_{\Omega_I\times\mathbb{R}^n} |v|^2|\nu|(dxdtdv) < \infty .$$

Define the subspaces $\mathbf{Z}$, $\mathbf{Z}_0$ of $\mathbf{C}$ as

$$\mathbf{Z}_0 := \left\{ z = \phi_t + v\cdot\nabla_x\phi \;\; ; \;\; \phi \in C_0^1(\Omega_I) \right\} \subset \mathbf{Z} := \left\{ z = \phi_t + v\cdot\nabla_x\phi \;\; ; \;\; \phi \in C^1(\Omega_{I_0}) \cap LIP(\Omega_I) \right\} .$$

Given $\mu_0, \mu_1 \in \mathcal{M}$, define h_{μ_0,μ_1} as a linear functional on $\mathbf{Z}$ as follows:

$$< h_{\mu_0,\mu_1}, z > := \int_\Omega \phi(x,T)\mu_1(dx) - \int_\Omega \phi(x,0)\mu_0(dx) \quad \text{for } z \in \mathbf{Z} , \tag{4.2}$$

(in particular, $< h_{\mu_0,\mu_1}, z >= 0$ if $z \in \mathbf{Z}_0$).

Lemma 4.1. *The functional h_{μ_0,μ_1}, so defined, is continuous (bounded) on $\mathbf{C}$.*

Proof. Let λ be a probability distribution on $\Omega \times \Omega$ so that $\pi^{(0)}_\#\lambda = \mu_0$, $\pi^{(1)}_\#\lambda = \mu_1$. Then

$$\int_\Omega \phi(x,T)\mu_1(dx) - \int_\Omega \phi(x,0)\mu_0(dx) = \int\int_{\Omega\times\Omega} [\phi(y,T) - \phi(x,0)]\,\lambda(dxdy) . \tag{4.3}$$

Now, for $\zeta(s) := \frac{(T-s)x+sy}{T}$ we obtain

$$\begin{aligned}\phi(y,T) - \phi(x,0) = \int_0^T \frac{d}{ds}\phi(\zeta(s),s)\,ds &= \int_0^T \left[\phi_t + \frac{y-x}{T}\cdot\nabla_x\phi \right]_{\zeta(s),s} ds \\ &= \int_0^T z\left(\zeta(s), s, \frac{y-x}{T}\right) ds \end{aligned} \tag{4.4}$$

In particular,

$$|\phi(y,T) - \phi(x,0)| \le \max_{(x,t)\in\Omega_I} \max_{|v|\le Diam(\Omega/T)} |z(x,t,v)| \le \|z\| \left[1 + \left(\frac{Diam(\Omega)}{T} \right)^2 \right] ,$$

where we used the definition on the norm $\|\cdot\|$ on $\mathbf{C}$ given by (4.1). The proof follows from (4.2,4.3) and since λ is a probability distribution on $\Omega \times \Omega$. □

The corresponding dual spaces are given by

$$\mathbf{Z}_0^* := \left\{ \nu \in \mathbf{C}^*; \int_{\Omega_I \times \mathbb{R}^n} z(x,t,v)\nu(dxdtdv) = 0 \ , \forall z \in \mathbf{Z}_0 \right\} \tag{4.5}$$

$$\supset \mathbf{Z}^*_{\mu_0,\mu_1} := \left\{ \nu \in \mathbf{C}^*; \int_{\Omega_I \times \mathbb{R}^n} z(x,t,v)\nu(dxdtdv) = < h_{\mu_0,\mu_1}, z > \ , \forall z \in \mathbf{Z} \right\} .$$

For any $\mu \in \mathbf{H}_2$, a convex subset of $\mathbf{C}^*$ is given by

$$\mathbf{C}^*_\mu := \left\{ \nu \in \mathbf{C}^* \ ; \int_{\Omega_I \times \mathbb{R}^n} \phi(x,t)\nu(dx,dt,dv) = \int_{\Omega_I} \phi(x,t)\mu(dxdt) \ \ \forall \ \phi \in C(\Omega_I) \right\} .$$

Finally, $F_\mu : \mathbf{C}^* \to \mathbb{R} \cup \{\infty\}$ is defined by

$$F_\mu(\nu) = \begin{cases} \frac{1}{2} \int_{\Omega_I \times \mathbb{R}^n} |v|^2 \nu(dxdtdv) & \text{if } \nu \in \mathbf{C}^*_\mu \\ \infty & \text{if } \nu \notin \mathbf{C}^*_\mu \end{cases}$$

We obtain

Lemma 4.2. *The function F_μ is convex on $\mathbf{C}^*$ for any $\mu \in \mathbf{H}_2$. In addition, if $F_\mu(\nu) < \infty$ and $\nu \in \mathbf{Z}_0^*$ then ν is a lifting of μ. Similarly, if $F_\mu(\nu) < \infty$ and $\nu \in \mathbf{Z}^*_{\mu_0,\mu_1}$ then $\mu \in \Lambda_2(\mu_0, \mu_1)$.*

Proof. The proof of Lemma 4.2 is almost evident from the definitions. Let us just prove the last part. Since $\mathbf{Z}^*_{\mu_0,\mu_1} \subset \mathbf{Z}_0^*$ it follows that ν is a lifting of $\mu \in \mathbf{H}_2$. We only have to show that $\mu \in \Lambda_2(\mu_0, \mu_1)$. Let $\phi \in C^1(\Omega_{I_0}) \cap LIP(\Omega_I)$, $\eta = \eta(t) \in C_0^1(I)$ satisfies $0 \leq \eta \leq 1$ on I and, for some $\varepsilon > 0$, $\eta(t) = 1$ for $\varepsilon \leq t \leq T - \varepsilon$, and $\eta_t \geq 0$ on $[0, \varepsilon]$, $\eta_t \leq 0$ on $[T - \varepsilon, T]$. Set $\phi^{(\varepsilon)} = \phi$ on $\Omega \times [\varepsilon, T - \varepsilon]$ and $\phi^{(\varepsilon)}(x,t) = \phi(x,\varepsilon)$ on $t \in [0, \varepsilon]$ (res. $\phi^{(\varepsilon)}(x,t) = \phi(x, T-\varepsilon)$ on $t \in [T - \varepsilon, T]$). Then $\eta\phi^{(\varepsilon)} \in C_0^1(\Omega_I)$, so

$$\begin{aligned} 0 = \int_{\Omega_I \times \mathbb{R}^n} \left[(\eta\phi^{(\varepsilon)})_t + \eta v \cdot \nabla_x \phi^{(\varepsilon)} \right] \nu(dxdtdv) = \int_\varepsilon^{T-\varepsilon} \int_{\Omega \times \mathbb{R}^n} \left[\phi_t + v \cdot \nabla_x \phi \right] \nu_{(t)}(dxdv)dt \\ + \int_0^\varepsilon \int_\Omega \eta_t \phi(x,\varepsilon)\mu_{(t)}(dx)dt + \int_{T-\varepsilon}^T \int_\Omega \eta_t \phi(x, T-\varepsilon)\mu_{(t)}(dx)dt \\ + \int_0^\varepsilon \int_{\Omega \times \mathbb{R}^n} \eta v \cdot \nabla_x \phi(x,\varepsilon)\nu_{(t)}(dxdv)dt + \int_{T-\varepsilon}^T \int_{\Omega \times \mathbb{R}^n} \eta v \cdot \nabla_x \phi(x, T-\varepsilon)\nu_{(t)}(dxdv)dt \ . \end{aligned} \tag{4.6}$$

Since ν is a lifting of some $\mu \in \mathbf{H}_2$ it follows that $\nu_{(t)}(dxdv)$ is a probability measure on $\Omega \times \mathbb{R}^n$. By the Cauchy-Schwartz inequality we estimate the last two integrals by $2\|\nabla_x\phi\|_\infty \sqrt{\mathbb{E}_\nu(|v|^2)}\varepsilon^{1/2}$. By Lemma 2.2, $\mu_{(t)}$ is Holder continuous of exponent $1/2$ in t, with respect to the W_1 topology, so

$$\begin{aligned} \int_0^\varepsilon \int_\Omega \eta_t \phi(x,\varepsilon)\mu_{(t)}(dx)dt = \int_0^\varepsilon \int_\Omega \eta_t \phi(x,\varepsilon)\mu_{(0)}(dx)dt + O(\varepsilon^{1/2})\|\nabla_x\phi\|_\infty \int_0^\varepsilon |\eta_t| dt \\ = \int_\Omega \phi(x,\varepsilon)\mu_{(0)}(dx) + O(\varepsilon^{1/2})\|\nabla_x\phi\|_\infty \int_0^\varepsilon \eta_t dt = \int_\Omega \phi(x,\varepsilon)\mu_{(0)}(dx) + O(\varepsilon^{1/2})\|\nabla_x\phi\|_\infty \ , \end{aligned} \tag{4.7}$$

using $\eta_t \geq 0$ on $[0,\varepsilon]$, hence $\int_0^\varepsilon |\eta_t| = \int_0^\varepsilon \eta_t = 1$. Similarly

$$\int_{T-\varepsilon}^{T} \int_\Omega (\eta\phi)_t \mu_{(t)}(dx)dt = -\int_\Omega \phi(x, T-\varepsilon)\mu_{(T)}(dx) + O(\varepsilon^{1/2})\|\nabla_x\phi\|_\infty . \tag{4.8}$$

Letting $\varepsilon \to 0$ we obtain from (4.6, 4.7,4.8):

$$\int_{\Omega_I \times \mathbb{R}^n} [\phi_t + v\cdot\nabla_x\phi]\,\nu(dxdtdv) - \int_\Omega [\phi(x,T)\mu_{(T)}(dx) - \phi(x,0)\mu_{(0)}(dx)] = 0 .$$

The above is valid for any $\phi \in C^1(\Omega_{I_0}) \cap LIP(\Omega_I)$. Since $\nu \in \mathbf{Z}^*_{\mu_0,\mu_1}$ by assumption, it follows that $\mu_{(0)} = \mu_0$ and $\mu_{(T)} = \mu_1$, hence $\mu \in \Lambda_2(\mu_0,\mu_1)$. □

Corollary 4.1. *If $\mu \in \mathbf{H}_2$ then*

$$\frac{1}{2}\|\mu\|_2^2 = -\inf_{\phi\in C_0^1}\left\{\int_{\Omega_I}(\phi_t + |\nabla_x\phi|^2/2)\mu(dxdt)\right\} = \frac{1}{2}\sup_{\phi\in C_0^1}\frac{\left(\int_{\Omega_I}\phi_t\mu(dxdt)\right)^2}{\int_{\Omega_I}|\nabla_x\phi|^2\mu(dxdt)} . \tag{4.9}$$

as well as

$$-\inf_{\phi\in C^1(\Omega_{I_0})\cap LIP(\Omega_I)}\left\{\int_{\Omega_I}(\phi_t+|\nabla_x\phi|^2/2)\mu(dxdt) + \int_\Omega\phi(x,0)\mu_0(dx) - \int_\Omega\phi(x,T)\mu_1(dx)\right\}$$

$$= \begin{cases} \frac{1}{2}\|\mu\|_2^2 & if\ \mu\in\Lambda_2(\mu_0,\mu_1) \\ \infty & if\ \mu\notin\Lambda_2(\mu_0,\mu_1) \end{cases} . \tag{4.10}$$

Proof. Certainly, F_μ satisfies all the conditions of Proposition 4.1. Using Lemma 4.2 and Proposition 4.1 in the definition of $\|\mu\|_2$ (Definition 2.1 for $p=2$) we obtain that

$$\frac{1}{2}\|\mu\|_2^2 = \inf_{\nu\in\mathbf{Z}^*} F_\mu(\nu) = \sup_{z\in\mathbf{Z}}\inf_{\nu\in\mathbf{C}^*}(F_\mu(\nu) - <\nu, z>)$$

$$= \sup_{\phi\in C_0^1(\Omega_I)}\inf_{\nu\in\mathbf{C}^*_\mu}\int_{\Omega_I\times\mathbb{R}^n}\left[\frac{1}{2}|v|^2 - \phi_t - v\cdot\nabla_x\phi\right]\nu(dxdtdv)$$

$$= \sup_{\phi\in C_0^1(\Omega_I)}\inf_{\nu\in\mathbf{C}^*_\mu}\int_{\Omega_I\times\mathbb{R}^n}\left[\frac{1}{2}|v-\nabla_x\phi|^2 - \phi_t - \frac{1}{2}|\nabla_x\phi|^2\right]\nu(dxdtdv)$$

$$= \sup_{\phi\in C_0^1(\Omega_I)}\left\{-\int_{\Omega_I}\left[\phi_t + \frac{1}{2}|\nabla_x\phi|^2\right]\mu(dxdt) + \frac{1}{2}\inf_{\nu\in\mathbf{C}^*_\mu}\int_{\Omega_I\times\mathbb{R}^n}|v-\nabla_x\phi|^2\nu(dxdtdv)\right\} .$$

So, we set $\nu = \mu\delta_{v-\nabla_x\phi}$ to annihilate the second integral and obtain the first equality in (4.9). For the second equality in (4.9) we observe

$$\inf_{\phi\in C_0^1}\left\{\int_{\Omega_I}(\phi_t+|\nabla_x\phi|^2/2)\mu(dxdt)\right\} = \inf_{\phi\in C_0^1}\inf_{\beta\in\mathbb{R}}\left\{\int_{\Omega_I}(\beta\phi_t+\beta^2|\nabla_x\phi|^2/2)\mu(dxdt)\right\}$$

$$= \inf_{\phi\in C_0^1}\left(-\frac{1}{2}\frac{\left(\int_{\Omega_I}\phi_t\mu(dxdt)\right)^2}{\int_{\Omega_I}|\nabla_x\phi|^2\mu(dxdt)}\right) .$$

Finally, we obtain (4.10) using the constraint $\mathbf{Z}^*_{\mu_0,\mu_1}$ for $\mathbf{Z}^*$ in Proposition 4.1. □

Example: Let $\mu = \sum_{i=1}^{k} \beta_k \delta_{(x-x_k(t))}$ where $x_k : I \to \Omega$ satisfies $\int_0^T |\dot{x}_j|^2 dt := |\dot{x}_j|_2 < \infty$ and $\beta_j \geq 0$, $\sum_j \beta_j = 1$. Then

$$\int_{\Omega_I} \phi_t \mu(dxdt) = \sum_j \beta_j \int_0^T \frac{\partial \phi}{\partial t}(x_j(t), t) dt$$

and

$$\int_{\Omega_I} |\nabla_x \phi|^2 \mu(dxdt) = \sum_j \beta_j \int_0^T |\nabla_x \phi|^2 (x_j(t), t) dt$$

On the other hand,

$$\int_0^T \frac{\partial \phi}{\partial t}(x_j(t), t) dt = \int_0^T \left[\frac{d\phi}{dt}(x_j(t), t) - \dot{x}_j(t) \cdot \nabla_x \phi(x_j(t), t) \right] dt$$
$$= - \int_0^T \dot{x}_j(t) \cdot \nabla_x \phi(x_j(t), t) dt$$

so, by an application (twice) of the Cauchy-Schwartz inequality,

$$\frac{\left(\int_{\Omega_I} \phi_t \mu(dxdt)\right)^2}{\int_{\Omega_I} |\nabla_x \phi|^2 \mu(dxdt)} = \frac{\left(\sum_j \beta_j \int_0^T \nabla_x \phi(x_j(t), t) \cdot \dot{x}_j dt\right)^2}{\sum_j \beta_j \int_0^T |\nabla_x \phi(x_j(t), t)|^2 dt} \leq \sum_j \beta_j \int_0^T |\dot{x}_j|^2 dt \ .$$

In fact, it can be shown that $||\mu||_2^2$ coincides with the above sum, and that there exists a maximizing sequence $\phi_n(x, t)$ such that $\nabla_x \phi_n(x_j(t), t) \to \dot{x}_j(t)$ for all j and a.e $t \in I$ (even if some of the orbits x_j intersect (!)-see [W]).

4.2 Dual representation of the Lagrangian

We shall now define a strong convex perturbation of the Lagrangian L_P (Definition 3.1). Let also $F : \mathbb{R} \to \mathbb{R}^+ \cup \{\infty\}$ such that

$$F(q) = \infty \ \text{ if } \ q < 0 \ \ ; \ \ F(0) = 0 \ \ ; \ \ cq^\omega < F(q) < Cq^\omega \text{ if } q > 0 \tag{4.11}$$

where $1 < \omega < 1 + 1/(n+1)$ and $c, C > 0$. The functional $\mathcal{I}_\varepsilon^P : \mathbf{C}^* \to \mathbb{R} \cup \{\infty\}$ is defined by:

$$\mathcal{I}_\varepsilon^P(\nu) := \int_{\Omega_I \times \mathbb{R}^n} \varepsilon F(f_\nu)\, dxdtdv + \frac{1}{2} \int_{\Omega_I \times \mathbb{R}^n} |v|^2 \nu(dxdtdv) + \int_{\Omega_I \times \mathbb{R}^n} P(x, t) \nu(dxdtdv), \tag{4.12}$$

if $\nu = f_\nu(x, t, v) dxdtdv$ is absolutely continuous with respect to Lebesgue measure and the density f_ν satisfies $F(f_\nu) \in \mathbb{L}^1(\Omega_I \times \mathbb{R}^n)$. Otherwise $\mathcal{I}_\varepsilon^P(\nu) = \infty$. Note that, since $F(q) = \infty$ for $q < 0$, it follows that $\mathcal{I}_\varepsilon^P(\nu) = \infty$ if $\nu \in \mathbf{C}^*$ is not a non-negative measure. However, $\mathcal{I}_\varepsilon^P$ can attain a finite value also for a measure ν which is not normalized (i.e not a probability measure on $\Omega_I \times \mathbb{R}^n$).

Given $\mu_0, \mu_1 \in \mathcal{M}$, define

$$I_\varepsilon^P(\mu_0, \mu_1) := \inf_{\nu \in \mathbf{Z}^*_{\mu_0, \mu_1}} \mathcal{I}_\varepsilon^P(\nu) \ . \tag{4.13}$$

Next, we claim

Lemma 4.3. *For any $\varepsilon > 0$,*

$$I^P_\varepsilon(\mu_0, \mu_1) \geq \mathcal{L}(\mu_0, \mu_1)$$

where $\mathcal{L}(\mu_0, \mu_1)$ as in Definition 3.1.

Proof. First, we can restrict ourselves to non-negative measures $\nu \in \mathbf{Z}^*_{\mu_0,\mu_1}$, since otherwise $\int F(f_\nu) = \infty$ by (4.11). We only have to show that if $\nu \geq 0$ and $\nu \in \mathbf{Z}^*_{\mu_0,\mu_1}$ then ν is a lifting of some $\mu \in \Lambda_2(\mu_0, \mu_1)$.

Using Lemma 4.2 it is, therefore, enough to prove that $\nu_t(dxdv)$ is a probability measure on $\Omega \times \mathbb{R}^n$ for a.e (Borel) $t \in I$. Setting $\phi(x,t) = \eta(t) \in C^1_0(I)$ we obtain from (4.5) that

$$\int_I \left(\int_{\Omega \times \mathbb{R}^n} \nu_t(dxdv) \right) \frac{d\eta}{dt} dt = 0$$

for any such η. This implies that $\int_{\Omega\times\mathbb{R}^n} \nu_t(dxdv)$ is constant for a.e. $t \in I$. Since $\nu \geq 0$ it implies that ν_t is a constant multiple of some probability measure on $\Omega \times \mathbb{R}^n$ for a.e. $t \in I$. This constant equals one since the Ω marginal of ν_t is C^* continuous on I by Lemma 2.2 and is a probability measure at $t = 0$ (μ_0) and $t = T$ (μ_1). □

We now proceed to a dual formulation of the constraint minimization of $\mathcal{I}^P_\varepsilon$. Certainly $\mathcal{I}^P_\varepsilon$ satisfies the assumption on $\mathcal{F}$ introduced in Proposition 4.1. In fact, it follows that the set $\{\nu \in \mathbf{C}^* \ ; \ \mathcal{I}^P_\varepsilon(\nu) < C\}$ is bounded (and hence $*$−compact) for any real C. Then Proposition 4.1 and (4.13) yield

$$I^P_\varepsilon(\mu_0, \mu_1) = \sup_{z \in \mathbf{Z}} \inf_{\nu \in \mathbf{C}^*} \left[\mathcal{I}^P_\varepsilon(\nu) - < \nu, z > + < h_{\mu_0,\mu_1}, z > \right]$$

$$= \sup_{\phi \in C^1(\Omega_I)} \inf_f \int_{\Omega_I \times R^n} \left[\varepsilon F(f) - f \left(\phi_t + v \cdot \nabla_x \phi - \frac{1}{2}|v|^2 - P \right) \right] dxdtdv +$$

$$\int_\Omega \phi(x,T)\mu_1(dx) - \phi(x,0)\mu_0(dx) \ ,$$

where $\inf_f$ stands for the infimum on all measurable functions on $\Omega_I \times \mathbb{R}^n$. Let

$$H_\varepsilon(f,\phi) := \int_{\Omega_I \times R^n} \left[\varepsilon F(f) - \left(\phi_t + v \cdot \nabla_x \phi - \frac{1}{2}|v|^2 - P \right) f \right] dxdtdv \ .$$

$$= \int_{\Omega_I \times R^n} dxdtdv \left[\varepsilon F(f) - \left(\phi_t + \frac{1}{2}|\nabla_x \phi|^2 - \frac{1}{2}|v - \nabla_x \phi|^2 - P \right) f \right]$$

Let F^* be the Legendre transform of F:

$$F^*(\lambda) = \sup_s [s\lambda - F(s)] \ .$$

By our assumption we know that F^* is also convex and non-negative on $\mathbb{R}$. It satisfies $F^*(\lambda) = 0$ for $\lambda \leq 0$. Now,

$$\inf_f H_\varepsilon(f,\phi) = -\varepsilon \int_{\Omega_I \times \mathbb{R}^n} F^* \left(\frac{\phi_t + |\nabla_x \phi|^2/2 - |v - \nabla_x \phi|^2/2 - P}{\varepsilon} \right) dxdtdv$$

$$= -\varepsilon^{1+n/2} \int_{\Omega_I \times \mathbb{R}^n} F^* \left(\frac{\phi_t + |\nabla_x \phi|^2/2 - P}{\varepsilon} - \frac{|v|^2}{2} \right) dx dt dv$$

Let

$$G(s) := \int_{\mathbb{R}^n} F^*(s - |v|^2/2) dv \tag{4.14}$$

and

$$\Psi_\varepsilon(\phi) := -\varepsilon^{1+n/2} \int_{\Omega_I} G\left(\frac{\phi_t + |\nabla_x \phi|^2/2 - P}{\varepsilon} \right) dx dt + \int_\Omega \phi(x, T)\mu_1(dx) - \phi(x, 0)\mu_0(dx) \tag{4.15}$$

We have proved:

Lemma 4.4. *For $\varepsilon > 0$ and $\mu_0, \mu_1 \in \mathcal{M}$,*

$$I_\varepsilon^P(\mu_0, \mu_1) = \sup_{\phi \in C^1(\Omega_{I_0}) \cap LIP(\Omega_I)} \Psi_\varepsilon(\phi) .$$

We shall also need the following result, whose proof is direct and omitted:

Lemma 4.5. *If F satisfies (4.11) then, for some constant $c > 0$, the function G defined in (4.14) satisfies $cq^{\omega/\omega-1} < G(q) < c^{-1}q^{\omega/\omega-1}$. Thus, the first integral of (4.15) is estimated by*

$$-\varepsilon^{1+n/2} \int_{\Omega_I} G\left(\frac{\phi_t + |\nabla_x \phi|^2/2 - P}{\varepsilon} \right) dx dt = -O(\varepsilon^{-\alpha}) \int_{\Omega_I} \left| \phi_t + \frac{|\nabla_x \phi|^2}{2} - P \right|^s dx dt$$

where $\alpha = 1/(\omega - 1) - n/2 > 0$ and $s = \omega/(\omega - 1) > 1 + n$ (c.f. (4.11)).

We also need:

Lemma 4.6. *Let $\mu_0, \mu_1 \in \mathcal{M}$. Then there exists a connecting orbit $\mu \in \Lambda_2(\mu_0, \mu_1)$ of finite $\mathbf{H}_2$ norm and a lifting ν such that both μ and ν has densities in $\mathbb{L}^p(\Omega_I)$ (res. $\mathbb{L}^p(\Omega_I \times \mathbb{R}^n)$), where $1 \leq p < 1 + 1/n$.*

In particular. it follows that for such ν as guaranteed in Lemma 4.6, each of the integrals in (4.12) is finite. Hence, there exists $C > 0$ (independent of ε) and $\nu \in \mathbf{Z}^*_{\mu_0,\mu_1}$ such that $\mathcal{I}_\varepsilon^P(\nu) < C$ for *any* $\varepsilon > 0$. In particular, $I_\varepsilon^P(\mu_0, \mu_1) < C$ for any such ε by (4.13). It follows from this, Lemma 4.3 and Lemma 4.4 that

Corollary 4.2. *For any $\mu_0, \mu_1 \in \mathcal{M}$ there exists $C > 0$ independent of ε where*

$$C > \sup_{\phi \in C^1(\Omega_{I_0}) \cap LIP(\Omega_I)} \Psi_\varepsilon(\phi) \geq \mathcal{L}(\mu_0, \mu_1) .$$

Lemma 4.6 is a direct result from Lemma 4.7 below. For its presentation we define the space $\mathbf{H}_p([t_0, t_1])$ by restricting $\mathbf{H}_p = \mathbf{H}_p(I)$ to orbits defined for a time interval $[t_0, t_1]$. The norm of $\mu \in \mathbf{H}_p([t_0, t_1])$ is denoted by $||\mu||_{2,[t_0,t_1]}$. Lemma 4.7 is also used in the proof of Lemma 5.8.

Lemma 4.7. *For any $t_1 > t_2 \in I$, any path $\overline{x} = \overline{x}(t) : [t_0, t_1] \to \Omega$ and any $\alpha > 0$ there exists an orbit $\mu_{(t)}dt \in \mathbf{H}_p([t_0, t_1])$ with $\mu_{(t_0)} = \delta_{x_0}$, $\mu_{(t_1)} = \delta_{x_1}$ where $x_i = \overline{x}(t_i)$, $i = 0, 1$, such that*

$$||\mu||^2_{2,[t_0,t_1]} \leq \int_{t_0}^{t_1} \left|\dot{\overline{x}}\right|^2 dt + C|t_1 - t_0|\alpha^{-2}$$

and $\mu_{(t)}(dx) = \rho(x,t)dx$ where $\rho \in \mathbb{L}^p(\Omega \times [t_0, t_1])$ for any $p \in [1, 1 + 1/n)$. Moreover

$$|\rho|_p \leq C(p) \left[(t_1 - t_0)\left(\frac{\alpha}{t_1 - t_0}\right)^{n(p-1)}\right]^{1/p}$$

and

$$supp(\rho) \subset \left\{(x,t) \in \Omega \times [t_0, t_1] \ ; \ |\overline{x}(t) - x| \leq C\frac{t_1 - t_0}{\alpha} \quad \forall\, t \in [t_0, t_1]\right\} .$$

In particular, the choice $\overline{x}(t) = x_0 + \frac{t-t_0}{t_1-t_0}(x_1 - x_0)$ yields

$$||\mu||^2_{2,[t_0,t_1]} \leq \frac{|x_1 - x_0|^2}{t_1 - t_0} + C|t_1 - t_0|\alpha^{-2} .$$

The proof of Lemma 4.7 is given in the Appendix.

5 Proof of main results

5.1 On the Hamilton-Jacobi Equation

In this section we introduce some fundamental results for the HJ equation

$$\phi_t + \frac{1}{2}|\nabla_x \phi|^2 = P \quad (x,t) \in \Omega_I \tag{5.1}$$

where $P \in C^1(\Omega_I)$. The book of L.Evans [E] contains a detailed exposition on the Hamilton-Jacobi equation. However, the discussion in [E] is restricted to generalized solutions of viscosity type and for time independent Hamiltonians, which excludes the application of backward solutions and time dependent pressure $P = P(x,t)$. The results in this section are all needed for the proof of the Main Theorem in section 3

We list below some properties of the action J_P (3.2):

Lemma 5.1. *For $P \in C^1(\Omega_I)$, the action J_P is satisfies the following:*

(a) For $\tau_1 < \tau_2 \in [0,T]$ and $x_1, x_2 \in \mathbb{R}^n$, the value of the action $J_P(x_1, x_2, \tau_1, \tau_2)$ is realized along a (possibly not unique) orbit $\overline{x}$ which satisfies the equation

$$\ddot{\overline{x}} = \nabla_x P(\overline{x}(t), t) \quad ; \quad t \in [\tau_1, \tau_2] \ . \tag{5.2}$$

(b) Assume further that there exists $C(t) > 0$ so that $P(x,t) - C(t)|x|^2$ is a concave function on $\mathbb{R}^n$ for any $t \in I_0$. Let $\overline{x}$ be an optimizer orbit connecting x_1, τ_1 to x_2, τ_2. For any $y \in \mathbb{R}^n$ and $t \in (0,T)$

$$J_P(x_1, x_2, \tau_1, \tau_2) - J_P(x_1, y, \tau_1, t) \geq$$

$$\dot{\bar{x}}(\tau_2)\cdot(x_2-y)+\left[P(x_2,\tau_2)-\frac{1}{2}|\dot{\bar{x}}|(\tau_2)\right](\tau_2-t)-O(|x_2-y|^2)-o(t-\tau_2)\ , \tag{5.3}$$

$$J_P(y,x_2,t,\tau_2)-J_P(x_1,x_2,\tau_1,\tau_2)\le$$

$$\dot{\bar{x}}(\tau_1)\cdot(x_1-y)+\left[P(x_1,\tau_1)-\frac{1}{2}|\dot{\bar{x}}|(\tau_1)\right](\tau_1-t)+O(|x_1-y|^2)+o(t-\tau_1)\ . \tag{5.4}$$

(c) For any $x_1, y, x_2 \in \mathbb{R}^n$, $t_1 < \tau < t_2$

$$J_P(x_1,y,t_1,\tau)+J_P(y,x_2,\tau,t_2)\ge J_P(x_1,x_2,t_1,t_2) \tag{5.5}$$

holds .

(d) For any pair $x_1, x_2 \in \mathbb{R}^n$ and a triple $t_1 < \tau < t_2$ there exists $y^* \in \mathbb{R}^n$ (possibly non-unique) for which the equality holds in (5.5)*:*

$$J_P(x_1,y^*,t_1,\tau)+J_P(y^*,x_2,\tau,t_2)=J_P(x_1,x_2,t_1,t_2)\ . \tag{5.6}$$

There exists a (possibly non-unique) optimal orbit $\bar{x}$ connecting (x_1,t_1) to (x_2,t_2) such that $\bar{x}(\tau)=y$. However, for any *such optimal orbit, $\dot{\bar{x}}(\tau)$ is determined* uniquely.

(e) For any $t > t_1$ $x \in \mathbb{R}^n$ and a.e $y \in \mathbb{R}^n$

$$\frac{\partial}{\partial t}J_P(x,y,t_1,t)+\frac{1}{2}\left|\nabla_y J_P(x,y,t_1,t)\right|^2=P(y,t)\ . \tag{5.7}$$

Definition 5.1. *$\phi(x,t)$ is a* forward *solution of* (5.1) *iff, for any $x \in \Omega$ and $t_1 > t_0 \in I$*

$$\phi(x,t_1)=\inf_{y\in\mathbb{R}^n}\left[J_P(y,x,t_0,t_1)+\phi(y,t_0)\right]\ . \tag{F}$$

Likewise, ϕ is a backward *solution iff*

$$\phi(x,t_0)=\sup_{y\in\mathbb{R}^n}\left[-J_P(x,y,t_0,t_1)+\phi(y,t_1)\right]\ . \tag{B}$$

Remark: *It follows, by the remark proceeding* (3.2)*, that the right sides of (F) (res. (B)) defines a function which is $\mathbb{Z}^n$ periodic on $\mathbb{R}^n$, namely defined on Ω, if $\phi(\cdot,t_0)$ (res. $\phi(\cdot,t_1)$) is a function on Ω.*

For the special case of zero-pressure Hamilton-Jacobi equation, the action is reduced to

$$J_0(x_1,x_2,t_1,t_2)=\frac{|x_2-x_1|^2}{2(t_2-t_1)}$$

and definition 5.1 reduces to the (original) Hopf-Lax formula:

Definition 5.2. *A forward solution of the pressureless Hamilton-Jacobi equation*

$$\phi_t + \frac{1}{2}|\nabla_x \phi|^2 = 0$$

satisfies, for any $t_1 > t_0$ and $x \in \Omega$

$$\phi(x,t_1) = \inf_{y \in \mathbb{R}^n} \left[\frac{1}{2} \frac{|x-y|^2}{t_1 - t_0} + \phi(y,t_0) \right] \quad , \tag{$\mathbf{F_0}$}$$

while a backward solution satisfies

$$\phi(x,t_0) = \sup_{y \in \mathbb{R}^n} \left[-\frac{1}{2} \frac{|x-y|^2}{t_1 - t_0} + \phi(y,t_1) \right] \quad . \tag{$\mathbf{B_0}$}$$

A forward (backward) solution can be constructed from an initial (end) data at $t = 0$ $(t = T)$ as follows.

Lemma 5.2. *For any* continuous *initial data $\phi(,0)$ on Ω and $P \in LIP(\Omega_I)$,*

$$\phi(x,t) = \inf_{y \in \mathbb{R}^n} [J_P(y,x,0,t) + \phi(y,0)] \tag{5.8}$$

is a forward solution and satisfies (5.1) *a.e. Moreover, $\phi \in LIP(\Omega \times (0,T])$ and*

$$\frac{|\phi(x,t) - \phi(y,t)|}{|x-y|} \leq \frac{C}{t} \tag{5.9}$$

where C is a constant independent *on $\phi(,0)$. Likewise, for any continuous end data $\phi(,T)$*

$$\phi(x,t) = \sup_{y \in \mathbb{R}^n} [-J_P(x,y,0,t) + \phi(y,1)]$$

is a backward solution and satisfies (5.1) *a.e., $\phi \in LIP(\Omega \times [0,T))$ and*

$$\frac{|\phi(x,t) - \phi(y,t)|}{|x-y|} \leq \frac{C}{T-t} \tag{5.10}$$

If, in either cases, the end data $\phi(,0)$ (res. $\phi(,T)$) is Lipschitz on Ω, then the corresponding forward (backward) solution is in $LIP(\Omega_I)$.

Next, we establish the connection between generalized sub-solutions, as defined in (3.4), and forward/backward solutions:

Lemma 5.3. *Both forward and backward solutions are generalized sub-solutions in the sense of* (3.4). *A forward (backward) solution is a maximal (minimal) generalized sub-solution in the following sense: If ψ is a generalized sub-solution and ϕ is a forward (backward) solution so that $\psi(x,\tau) = \phi(x,\tau)$ for all $x \in \Omega$ and some $\tau \in I$, then $\phi(x,t) \geq \psi(x,t)$ ($\phi(x,t) \leq \psi(x,t)$) for all $x \in \Omega$ and $t \geq \tau$ ($t \leq \tau$) in I.*

An immediate corollary from Lemma 5.3 is:

Corollary 5.1. *Let ϕ be a forward solution and ψ is a backward solution on Ω_I.*

i) If $\phi(x,T) = \psi(x,T)$ holds $\forall x \in \Omega$ then $\psi(x,t) \leq \phi(x,t)$ $\forall (x,t) \in \Omega_I$.

ii) Similarly, if $\phi(x,0) = \psi(x,0)$ holds $\forall x \in \Omega$ then $\psi(x,t) \leq \phi(x,t)$ $\forall (x,t) \in \Omega_I$.

Next, we wish to address the notion of a *reversible* solution:

Definition 5.3. *A reversible pair $\{\overline{\phi}, \underline{\phi}\}$ where $\overline{\phi}$ ($\underline{\phi}$) is a forward (backward) solution on Ω_I such that $\overline{\phi}(x,0) = \underline{\phi}(x,0)$ and $\overline{\phi}(x,T) = \underline{\phi}(x,T)$ for any $x \in \Omega$. By Corollary 5.1, $\overline{\phi} \geq \underline{\phi}$ on Ω_I. For any such reversible pair we denote the reversibility set of the pair as the relatively closed set $K_0(\overline{\phi}, \underline{\phi}) \subset \Omega_{I_0}$ given by*

$$K_0(\overline{\phi}, \underline{\phi}) := \{(x,t) \in \Omega_{I_0} \ ; \ \ \overline{\phi}(x,t) = \underline{\phi}(x,t)\} \subset \Omega_{I_0}$$

Likewise,

$$K_0^{(t)}(\overline{\phi}, \underline{\phi}) = K_0(\overline{\phi}, \underline{\phi}) \cap [\Omega \times \{t\}] \quad \textit{for any} \quad t \in (0,T) \ .$$

If $\overline{\phi} \equiv \underline{\phi}$ then $\phi := \overline{\phi} = \underline{\phi}$ is called a reversible *solution.*

From Corollary 5.1 we obtain a way to create reversible pairs. It turns out that, in the case $P \equiv 0$, this way yields reversible solutions:

Lemma 5.4. *Given $\phi_0 \in LIP(\Omega)$, let ϕ be the forward solution subjected to $\phi(x,0) = \phi_0(x)$. Let $\underline{\psi}$ be the backward solution subjected to $\underline{\psi}(x,T) = \phi(x,T)$, and $\overline{\psi}$ the forward solution subjected to*

$$\overline{\psi}(x,0) = \underline{\psi}(x,0) \ .$$

Then $\{\overline{\psi}, \underline{\psi}\}$ is a reversible pair. Moreover, if $P \equiv 0$ then $\overline{\psi} = \underline{\psi}$ is a reversible solution.

The next Lemmas indicate that reversible pairs (in particular, reversible solutions) are closely related to *classical* solutions of the Hamilton-Jacobi equation.

Lemma 5.5. *If $\phi \in C^1(\Omega_I)$ is a classical solution of* (5.1) *then ϕ is a reversible solution.*

Using Lemma 5.1 we show that the converse of Lemma 5.5 also holds, in some sense:

Lemma 5.6. *If $\{\overline{\phi}, \underline{\phi}\}$ is a reversible pair then both $\overline{\phi}$ and $\underline{\phi}$ are differentiable on $K_0 := K_0(\overline{\phi}, \underline{\phi})$ (cf., Definition 5.3). Moreover, $\nabla\overline{\phi} := \{\nabla_x\overline{\phi}, \overline{\phi}_t\} = \nabla\underline{\phi} := \{\nabla_x\underline{\phi}, \underline{\phi}_t\}$ and the H.J equation is satisfied on this set. If, in addition, P satisfies the condition of Lemma 5.1-(b) then $\nabla\phi$ is locally Lipschitz continuous on K_0 and ϕ satisfies* (5.1) *pointwise on this set.*

Lemma 5.7. *Assume P satisfies the condition of Lemma 5.1-(b). Let $\boldsymbol{v}(x,t)$ be a Lipschitz extension of $\nabla_x\phi$ from K_0 to Ω_{I_0}. Then the set K_0 is invariant with respect to the (unique) flow generated by the vectorfield $\boldsymbol{v}$.*

Finally, we introduce the two following results, to be needed in Section 5.2:

Lemma 5.8. *If $\phi \in C^1(\Omega \times [t_0, t_1])$ satisfying*

$$\phi_t + \frac{1}{2}|\nabla_x \phi|^2 = P + \xi \quad ; \quad (x,t) \in \Omega \times [t_0, t_1]$$

where $P, \xi \in Lip(\Omega \times [t_0, t_1])$, $s > n+1$ and $||\xi||_s$ stands for the $\mathbb{L}^s(\Omega \times [t_0, t_1])$ norm of ξ then, for any x_0, x_1 in Ω, any $t_1 > t_0$ and any orbit $\overline{x} = \overline{x}(t) : [t_0, t_1] \to \Omega$ satisfying $\overline{x}(t_0) = x_0$, $\overline{x}(t_1) = x_1$:

$$\phi(x_1, t_1) - \phi(x_0, t_0) \leq \frac{1}{2}\int_{t_0}^{t_1} |\dot{\overline{x}}|^2 + \int_{t_0}^{t_1} P(\overline{x}(t), t)dt + C_1||\xi||_s^{2\beta}(t_1 - t_0)^{\lambda} + C_2||P||_{lip}(t_1 - t_0)^{\eta}||\xi||_s^{\beta},$$

where $\beta = \frac{p}{2p+n(p-1)}$, , $p = s* := \frac{s-1}{s}$, $\lambda = \frac{2+n-np}{2p+np-n}$, $\eta = \frac{4p+(n-1)(p-1)}{2p+n(p-1)}$.

From Lemma 5.8 and Definition 5.1 we also obtain

Corollary 5.2. *Let $\phi \in C^1(\Omega_I)$ be a solution and ψ a forward solution of the respective equations on Ω_I:*

$$\phi_t + 1/2|\nabla_x \phi|^2 = P + \xi \quad ; \quad \psi_t + 1/2|\nabla_x \psi|^2 = P$$

such that $\psi(x, 0) = \phi(x, 0)$ on Ω. Then

$$\psi(x, T) \geq \phi(x, T) - \left[C_1||\xi||_s^{2\beta} + C_2||P||_{lip}||\xi||_s^{\beta}\right]$$

where s, β as defined in Lemma 5.8.

5.2 Proof of the main Theorem

For the proofs of the results in Section 5.1 see [W1].

First, the existence of a minimizer for $\mathcal{L}(\mu_0, \mu_1)$ in $\Lambda_2(\mu_0, \mu_1)$ follows immediately by the lower-semi-continuity of $||\mu||_2$ and the local compactness of $\mathbf{H}_2$. Next, we shall prove the chain of inequalities:

$$\mathcal{E}(\mu_0, \mu_1) \geq \mathcal{L}(\mu_0, \mu_1) \geq \mathcal{K}(\mu_0, \mu_1) \geq \mathcal{E}(\mu_0, \mu_1)$$

from left to right, together with the existence of a maximizer for $\mathcal{E}(\mu_0, \mu_1)$ in $\overline{\Lambda}^*_P$.

- $\mathcal{E}(\mu_0, \mu_1) \geq \mathcal{L}(\mu_0, \mu_1)$

  From Lemma 4.5 and Corollary 4.2 there exists a sequence $\varepsilon_k \to 0$ and $\phi_k \in C^1(\Omega_I)$ such that

$$\mathcal{L} \leq \Psi_{\varepsilon_k}(\phi_k) < -O\left(\varepsilon_k^{-\alpha}\right)||\xi_k||_s^s + \int_\Omega \phi_k(x, T)\mu_1(dx) - \int_\Omega \phi_k(x, 0)\mu_0(dx) < C \quad (5.11)$$

  where $\alpha > 0$, $s > n + 1$ and

$$\xi_k = \phi_{k,t} + |\nabla_x \phi_k|^2/2 - P \quad .$$

  Let now $\Xi : \Omega_I \to \Omega$ be a flow such that

i) $\sup_{x\in\Omega} \int_0^T \left|\frac{\partial \Xi(x,t)}{\partial t}\right|^2 dt := E < \infty$,

ii) $\Xi(x,0) = x$, $\Xi_\#(\cdot,T)\mu_0 = \mu_1$

By Lemma 5.8 and (i) we obtain

$$\phi_k(\Xi(x,T),T) - \phi_k(x,0) \leq \frac{1}{2}E + |P|_\infty + C_1||\xi_k||_s^{2\beta} + C_2||P||_{lip}||\xi_k||_s^\beta \ .$$

Integrate the above against μ_0 on Ω and use (ii) to obtain

$$\int_\Omega \phi_k(x,T)\mu_1(dx) - \int_\Omega \phi_k(x,0)\mu_0(dx) \leq \frac{1}{2}E + |P|_\infty + C_1||\xi_k||_s^{2\beta} + C_2||P||_{lip}||\xi_k||_s^\beta \tag{5.12}$$

Using $s > n+1$ and $\beta < 1/2$ (c.f. Lemma 5.8) we obtain from (5.11) and (5.12) that $||\xi_k||_s \to 0$ as $\varepsilon_k \to 0$. In addition

$$\liminf_{k\to\infty} \left[\int_\Omega \phi_k(x,T)\mu_1(dx) - \int_\Omega \phi_k(x,0)\mu_0(dx)\right] \geq \mathcal{L} \ . \tag{5.13}$$

Now, we may replace the sequence ϕ_k by a sequence of forward solutions $\overline{\psi}_k$ of the equation

$$\overline{\psi}_{k,t} + \frac{1}{2}\left|\nabla_x \overline{\psi}_k\right|^2 = P \quad ; \quad \overline{\psi}_k(x,0) = \phi_k(x,0) \ .$$

This is also a maximizing sequence which, by Corollary 5.2 together with $||\xi_k||_s \to 0$, yields

$$\liminf_{k\to\infty} \left[\int_\Omega \overline{\psi}_k(x,T)\mu_1(dx) - \int_\Omega \overline{\psi}_k(x,0)\mu_0(dx)\right] \geq \mathcal{L} \ . \tag{5.14}$$

From Lemma 5.2 we also obtain a uniform estimate on $\overline{\psi}_k$ in $LIP(\Omega \times (0,T])$. In particular, the sequence $\overline{\psi}_k(;T)$ is uniformly Lipschitz on Ω.

Now, define $\underline{\psi}_k$ to be the *backward* solutions of (3.6) subjected to $\underline{\psi}_k(x,T) = \overline{\psi}_k(x,T)$. From the first part of Lemma 5.2, $\underline{\psi}_k(x,0) \leq \overline{\psi}_k(x,0)$ on Ω so (5.14) is satisfied for $\underline{\psi}_k$ as well. Moreover, by the last part of Lemma 5.2 $\underline{\psi}_k$ are uniformly bounded in the Lipschitz norm on Ω_I. So, we can extract a subsequence of $\underline{\psi}_k$ which converges uniformly on $LIP(\Omega \times [0,T])$ to a backward solution $\underline{\psi}$.In particular, *both* $\underline{\psi}(,T)$ and $\underline{\psi}(,0)$ are Lipschitz. Let $\overline{\psi}$ be the forward solution satisfying $\overline{\psi}(,0) = \underline{\psi}(,0)$. By Corollary 5.1 and definition 5.3, the pair $(\overline{\psi}, \underline{\psi})$ is a reversible pair and both functions are in $\overline{\Lambda}_P^*$ (see the first part of Lemma 5.3). Moreover, the inequality (5.14) is preserved in the limit process, so

$$\int_\Omega \psi(x,T)\mu_1(dx) - \int_\Omega \psi(x,0)\mu_0(dx) \geq \mathcal{L}$$

holds for both $\psi = \overline{\psi}$ and $\psi = \underline{\psi}$ (recall $\underline{\psi} = \overline{\psi}$ on $\Omega \times \{0\}$ and $\Omega \times \{T\}$).

- $\mathcal{L}(\mu_0,\mu_1)\geq\mathcal{K}(\mu_0,\mu_1)$
Recall that there exists a minimizer of $\mathcal{L}(\mu_0,\mu_1)$ by the first part of the Theorem. Let μ be such a minimizer. We now use the regularization Lemma 2.3 to approximate μ by smooth densities $\mu_n=\rho_n(x,t)dxdt$. Let $\boldsymbol{v}_n$ be the regularized velocity field. Then

$$\lim_{n\to\infty}\int_{\Omega_I}|\boldsymbol{v}_n|^2(x,t)\rho_n(x,t)dxdt=||\mu||_2^2\ . \tag{5.15}$$

as well as

$$\lim_{n\to\infty}\int_{\Omega_I}\rho_n(x,t)P(x,t)dxdt=\int_{\Omega_I}P\mu(dxdt)\ .$$

Define

$$m_n(x,t)=\rho_n(x,t)\boldsymbol{v}_n(x,t)\ . \tag{5.16}$$

Then $m_n\in C^1(\Omega_I)$. Define now

$$\boldsymbol{v}_n^\varepsilon(x,t)=\frac{m_n(x,t)}{\rho_n(x,t)+\varepsilon}$$

By assumption, $\boldsymbol{v}_n^\varepsilon$ is Lipschitz on Ω_I, $t\in I$. Define $\rho_n^{(\varepsilon)}(x,t)$ as the solution of

$$\frac{\partial\rho_n^{(\varepsilon)}}{\partial t}+\nabla_x\left[\boldsymbol{v}_n^\varepsilon\rho_n^{(\varepsilon)}\right]=0\ \ ;\ \ \rho_n^{(\varepsilon)}(x,0)=\rho_n(x,0)\ . \tag{5.17}$$

Since $\boldsymbol{v}_n^\varepsilon$ is Lipschitz, we may define the flow associated with it as $\Gamma^t_{(\varepsilon)}:\Omega\to\Omega$ for $t\in I$, namely $\Gamma^t_{(\varepsilon)}(x)=y_{(x)}(t)$ where $\dot{y}_{(x)}=\boldsymbol{v}_n^\varepsilon(y_{(x)}(t),t)$ and $y_{(x)}(0)=x$. It follows that $\Gamma^t_{(\varepsilon),\#}\rho_n(\cdot,0)dx=\rho_n^{(\varepsilon)}(\cdot,t)dx$ for all $t\in I$. In particular:

$$\mathcal{K}\left(\rho_n(x,0)dx,\rho_n^{(\varepsilon)}(x,T)dx\right)\leq\int_\Omega\rho_n(x,0)J_P\left(x,\Gamma^T_{(\varepsilon)}(x)0,T\right)dx$$

$$=\int_{\Omega_I}\rho_n(x,0)\frac{d}{dt}J_P\left(x,\Gamma^t_\varepsilon(x),0,t\right)dtdx$$

$$=\int_{\Omega_I}\rho_n(x,0)\left[\frac{\partial}{\partial t}J_P(x,\Gamma^t_{(\varepsilon)}(x),t)+\nabla_{y=\Gamma^t_\varepsilon(x)}J_P(x,y,0,t)\cdot\boldsymbol{v}_n^\varepsilon\left(\Gamma^t_\varepsilon(x),t\right)\right]dxdt \tag{5.18}$$

From (5.7)

$$\partial_tJ_P(x,y,0,t)+\nabla_yJ_P(x,y,0,t)\cdot\boldsymbol{v}_n^\varepsilon(y,t)=P(y,t)+\frac{1}{2}\left|\boldsymbol{v}_n^\varepsilon\right|^2(y,t)-\frac{1}{2}\left|\nabla_yJ_P(x,y,0,t)-\boldsymbol{v}_n^\varepsilon(y,t)\right|^2\ .$$

Substitute the above in (5.18) at $y=\Gamma^t_\varepsilon(x)$ to obtain

$$\mathcal{K}\left(\rho_n(x,0)dx,\rho_n^{(\varepsilon)}(x,T)dx\right)\leq\int_{\Omega_I}\rho_n(x,0)\left[P\left(\Gamma^t_\varepsilon(x),t\right)+\frac{1}{2}\left|\boldsymbol{v}_\varepsilon^n\right|^2\left(\Gamma^t_\varepsilon(x),t\right)\right]dxdt$$

$$=\int_{\Omega_I}\rho_n^{(\varepsilon)}(x,t)\left[\frac{1}{2}\left|\boldsymbol{v}_\varepsilon^n\right|^2(x,t)+P(x,t)\right]dxdt\leq\int_{\Omega_I}\rho_n^{(\varepsilon)}(x,t)\left[\frac{1}{2}\left|\boldsymbol{v}^n\right|^2(x,t)+P(x,t)\right]dxdt \tag{5.19}$$

where the last inequality follows from (5.16). We next show that

$$\lim_{\varepsilon\to 0}\int_\Omega \left|\rho_n^{(\varepsilon)}(x,t)-\rho_n(x,t)\right| dx = 0 \tag{5.20}$$

for any $t\in I$. In fact, we note that $\rho_n+\varepsilon$ solves equation (5.17), hence $w_n^{(\varepsilon)} := \rho_n-\rho_n^{(\varepsilon)}+\varepsilon$ solves this equation as well. Since $w_n^{(\varepsilon)}(x,0)=\varepsilon>0$ we obtain that $w_n^{(\varepsilon)}\geq 0$ over Ω_I and, moreover,

$$\int_\Omega \left|\rho_n(x,t)-\rho_n^{(\varepsilon)}(x,t)\right| dx - |\Omega|\varepsilon \leq \int_\Omega \left|w_n^{(\varepsilon)}(x,t)\right| = \int_\Omega \left|w_n^{(\varepsilon)}(x,0)\right| = |\Omega|\varepsilon$$

for all $t\in I$. Now we take first the limit $\varepsilon\to 0$ then the limit $n\to\infty$ in (5.19). The r.h.s of (5.19) converges to $\mathcal{L}(\mu_0,\mu_1)$. Now, $\rho_n(x,0)dx$ and $\rho_n^\varepsilon(x,T)dx$ converges, as $n\to\infty$ and $\varepsilon\to 0$, weak$-*$ to μ_0 and μ_1, respectively. Since $\mathcal{K}$ is lower-semi-continuous in both μ_0 and μ_1 we obtain the desired result from (5.19) .

- $\mathcal{E}(\mu_0,\mu_1)\leq\mathcal{K}(\mu_0,\mu_1)$.
  Let $\lambda\in C^*(\Omega\times\Omega)$ be an optimizer of $\mathcal{K}$. Since $\pi_\#^{(1)}\lambda=\mu_1$ then

$$\int_\Omega \phi_1(x)\mu_1(dx) = \int_\Omega\int_\Omega \phi_1(y)\lambda(dxdy) \quad\text{and}\quad \int_\Omega \phi_0(x)\mu_0(dx) = \int_\Omega\int_\Omega \phi_0(x)\lambda(dxdy)$$

  for any continuous ϕ_1,ϕ_2. Set $\phi_1(x)=\psi(x,T)$ and $\phi_0(x)=\psi(x,0)$ with ψ an optimal backward solution of problem $\mathcal{E}$. Then

$$\mathcal{E} = \int_\Omega \psi(x,T)\mu_1(dx) - \int_\Omega \psi(x,0)\mu_0(dx) = \int_\Omega\int_\Omega \left[\psi(y,T)-\psi(x,0)\right]\lambda(dxdy)\ .$$

  Since ψ is a backward solution then

$$\int_\Omega\int_\Omega \left[\psi(y,T)-\psi(x,0)\right]\lambda(dxdy) \leq \int_\Omega\int_\Omega J_P(x,y,0,T)\lambda(dxdy) = \mathcal{K}\ .$$

We have proved

$$\int_\Omega \overline{\psi}(x,T)\mu_1(dx) - \int_\Omega \overline{\psi}(x,0)\mu_0(dx) = \int_\Omega \underline{\psi}(x,T)\mu_1(dx) - \int_\Omega \underline{\psi}(x,0)\mu_0(dx) = \mathcal{L}(\mu_0,\mu_1)\ . \tag{5.21}$$

We now turn to the proof of parts (i)-(vi) of the Theorem.

i) Let $\mu^{(0)}$ be a minimizer of $\mathcal{L}$. Given $t\in I_0$, let $\mu_{1/2}:=\mu^{(0)}_{(t)}\in\mathcal{M}$. Let us consider $\mu^{(1)}$ to be the restriction of $\mu^{(0)}$ to $\Omega\times[0,t]$ and $\mu^{(2)}$ the restriction of $\mu^{(0)}$ to $\Omega\times[t,T]$. Evidently, $\mu^{(1)}$ is a minimizer of L_P on the set of orbits $\Lambda_2(\mu_0,\mu_{1/2})$ confined to $[0,t]$ while $\mu^{(2)}$ is a minimizer on $\Lambda_2(\mu_{1/2},\mu_1)$ with respect to the same set, confined to $[t,T]$. In particular,

$$L_P(\mu^{(1)})+L_P(\mu^{(2)}) = L_P(\mu^{(0)}) = \mathcal{L}(\mu_0,\mu_1)\ . \tag{5.22}$$

By what we know so far,

$$\int_\Omega \overline{\psi}(x,t)\mu_{1/2}(dx) - \int_\Omega \overline{\psi}(x,0)\mu_0(dx) \le L_P(\mu^{(1)}) \tag{5.23}$$

$$\int_\Omega \overline{\psi}(x,T)\mu_1(dx) - \int_\Omega \overline{\psi}(x,t)\mu_{1/2}(dx) \le L_P(\mu^{(2)}) . \tag{5.24}$$

However, if we sum (5.23) and (5.24) and use (5.22) and (5.21), we conclude that there is, in fact, an equality in both (5.23) and (5.24). Same argument holds for $\underline{\psi}$ as well. Thus

$$\int_\Omega \overline{\psi}(x,t)\mu_{1/2}(dx) - \int_\Omega \overline{\psi}(x,0)\mu_0(dx) = L_P(\mu^{(1)}) = \int_\Omega \underline{\psi}(x,t)\mu_{1/2}(dx) - \int_\Omega \underline{\psi}(x,0)\mu_0(dx) .$$

Since $\overline{\psi}(x,0) \equiv \underline{\psi}(x,0)$,

$$\int_\Omega \left[\overline{\psi}(x,t) - \underline{\psi}(x,t)\right] \mu_{1/2}(dx) = 0 .$$

But, $\overline{\psi} \ge \underline{\psi}$ by Lemma 5.3. Hence $\overline{\psi}(x,t) = \underline{\psi}(x,t)$ on $supp\left(\mu^{(0)}_{(t)}\right) = supp\left(\mu_{1/2}\right)$. This, together with Lemma 5.6, proves that $Supp\left(\mu^{(0)}\right) \cap \Omega_{I_0} \subset K_0 \equiv \{(x,t) \ ; \ t \in I_0 \ , \underline{\psi}(x,t) = \overline{\psi}(x,t)\}$ and, in particular, that ϕ is differentiable at *any* point on the support of $\mu^{(0)}$ in Ω_I.

ii) This part follows from Lemma 5.7. In addition, the limits $\lim_{\tau \to T} \mathbf{T}^\tau_t$ and $\lim_{\tau \to 0} \mathbf{T}^t_\tau$ exists since $\nabla_x \psi$ is uniformly bounded on K_0. In particular, the Lipschitz extension $\boldsymbol{v}$ can be chosen to be a uniformly bounded function on Ω_I as well.

iii) Suppose there are two optimal solutions ψ_1, ψ_2 of $\mathcal{E}(\mu_0, \mu_1)$. To prove the uniqueness for the vector field $\boldsymbol{v}$ we claim that

$$\int_{\Omega_I} \left|\nabla_x \psi_1 - \nabla_x \psi_2\right|^2 \mu(dxdt) = 0$$

for any minimizer $\mu \in \Lambda_2(\mu_0, \mu_1)$ of $\mathcal{L}(\mu_0, \mu_1)$. Let $\psi = \alpha\psi_1 + (1-\alpha)\psi_2$ where $\alpha \in (0,1)$. Then

$$|\nabla_x \psi|^2 = \alpha|\nabla_x \psi_1|^2 + (1-\alpha)|\nabla_x \psi_2|^2 - \alpha(1-\alpha)|\nabla_x \psi_1 - \nabla_x \psi_2|^2 ,$$

so $\psi_t + |\nabla_x \psi|^2/2 < P$ and

$$\int_{\Omega_I} \left[\psi_t + |\nabla_x \psi|^2/2\right] \mu(dxdt) < \int_{\Omega_I} P\mu(dxdt) \tag{5.25}$$

if $\nabla_x \psi_1 \ne \nabla_x \psi_2$ at *some* point in the support of a minimizer μ (recall that both $\nabla \psi_i$, $i = 1,2$ are continuous on the support of ψ by Lemma 5.6).

On the other hand,

$$\mathcal{E}(\mu_0, \mu_1) = \int_\Omega \left[\psi(x,T)\mu_1(dx) - \psi(x,0)\mu_0(dx)\right] = \mathcal{L}(\mu_0, \mu_1) \tag{5.26}$$

follows from the assumptions that both ψ_1, ψ_2 are maximizers of $\mathcal{E}$. From (4.10), (5.25) and (5.26) it follows that

$$L_P(\mu) := \frac{1}{2}||\mu||_2^2 + \int_{\Omega_I} P\mu(dxdt) > \frac{1}{2}||\mu||_2^2 + \int_{\Omega_I} (\psi_t + |\nabla_x\psi|^2/2)\mu(dxdt) \geq \mathcal{E}(\mu_0, \mu_1) = \mathcal{L}(\mu_0, \mu_1) ,$$

in contradiction to the assumption that μ is a minimizer of L_P.

iv) Let, again, $\mu \in \Lambda_2(\mu_0, \mu_1)$ a minimizer of $\mathcal{L}$ and ψ a maximizer of $\mathcal{E}$. Since ψ satisfies the HJ equation on a closed set K_0 containing the support of μ in Ω_{I_0} and is a C^1 function there, we can extend it as a C^1 function on Ω_{I_0} so $\psi \in C^1(\Omega_{I_0}) \cap LIP(\Omega_I)$ and, by (5.26),

$$-\int_{\Omega_I} \left[\psi_t + \frac{1}{2}|\nabla_x\psi|^2 - P\right] \mu(dxdt) + \int_{\Omega} [\psi(x,T)\mu_1(dx) - \psi(x,0)\mu_0(dx)] = \mathcal{L}(\mu_0, \mu_1) = L_P(\mu) \tag{5.27}$$

We now use Corollary 4.1 (4.10) to observe that ψ is a maximizer of the left of (5.27), so by taking the variation $\phi = \psi + \varepsilon\eta$ with $\eta \in C^1(\Omega_I)$ we obtain

$$\int_{\Omega_I} (\eta_t + \nabla_x\psi \cdot \nabla_x\eta)\mu(dxdt) + \int_{\Omega} \eta(x,0)\mu_0(dx) - \int_{\Omega} \eta(x,T)\mu_1(dx) \geq 0$$

for *any* such η. Replacing η by $-\eta$ we obtain the equality above. Moreover, by the same argument following (5.22) to (5.24) we also obtain that

$$\int_{t_0}^{t} \int_{\Omega} (\eta_t + \nabla_x\psi \cdot \nabla_x\eta)\mu_t(dx) + \int_{\Omega} \eta(x,t_0)\mu_{(t_0)}(dx) - \int_{\Omega} \eta(x,t)\mu_{(t)}(dx) = 0 \ , \tag{5.28}$$

hold for any $0 < t_0 < t < T$. In particular, μ solves the weak form of the continuity equation with $\boldsymbol{v} = \nabla_x\psi$.

Now, we know that, by the additional assumption on P, that K_0 is invariant with respect to the flow $\mathbf{T}_{t_0}^t$ induced by the *Lipschitz* vectorfield $\boldsymbol{v}$ extending $\nabla_x\psi$. We shall now prove that μ is transported by this flow. That is, for any choice of $t_0, t \in (0,T)$, we need to show that $\mu_{(t)} = \gamma_t$ where

$$\gamma_t := \left[\mathbf{T}_{t_0}^t\right]_{\#} \mu_{t_0} \ .$$

Since $\mathbf{T}$ is the flow generated by $\boldsymbol{v}$ and K_0 is invariant with respect to $\boldsymbol{v}$ it follows that $\gamma = \gamma_t dt$ is supported on K and solves the weak form of the continuity equation as well. Setting $\zeta_t = \mu_{(t)} - \gamma_t$, $\zeta := \zeta_t dt$ we obtain from (5.28)

$$\int_{t_0}^{t} \int_{\Omega} (\eta_\tau + \boldsymbol{v} \cdot \nabla_x\eta)\zeta_\tau(dx)d\tau = \int_{\Omega} \eta(x,t)\zeta_t(dx) \tag{5.29}$$

for any $\eta \in C^1([t_0, t]; \mathbb{R})$ where we used $\zeta_{t_0} \equiv 0$.

Let now $h = h(x) \in C^1(\Omega)$. Let $\eta = \eta(x,\tau)$ be a solution of

$$\eta_\tau + \boldsymbol{v} \cdot \nabla_x\eta = 0 \ \ ; \ \ \eta(x,t) = h(x) \ , \ t_0 \leq \tau \leq t \ . \tag{5.30}$$

Since, by Lemma 5.6 and Lemma 5.7, the vector field $\nabla_x \psi = \boldsymbol{v}$ is locally Lipschitz continuous on K_0 which is invariant with respect to the induced flow, we can find a solution of (5.30) on $K \cap (\Omega \times [t_0, t])$ via

$$\eta(x, \tau) = h\left(T_\tau^t(x)\right) . \tag{5.31}$$

The function η so defined can be extended into a C^1 function on $\Omega \times [t_0, t]$. It satisfies (5.30) on K_0, so, recalling that ζ is supported on K_0, we substitute now (5.30) in (5.29) to obtain $\zeta_t \equiv 0$ and the proof of part (iv).

v) The optimality of $\mathbf{T}$ is evident from the proof of (iii) and (iv).

vi) From the last part of Lemma 5.4 it follows that ψ is a reversible solution so Lemma 5.6 implies that $\psi_t + |\nabla_x \psi|^2/2 = 0$ is satisfied *everywhere* on Ω_{I_0}. The flow induced by such a solution is given by $\mathbf{T}_\tau^t(x) = x + (t-\tau)\nabla_x \psi(x, \tau)$ and, by (iv) and (v), it transports $\mu_{(\tau)}$ to $\mu_{(t)}$ optimally.

6 Appendix

Proof. of Proposition 4.1:
Define

$$\Phi(c^*, z) := \mathcal{F}(c^*) - < c^*, z > + < h, z > .$$

First, note that

$$I = \inf_{c^* \in \mathbf{C}^*} \sup_{z \in \mathbf{Z}} \Phi(c^*, z) .$$

Indeed, if $c^* \notin \mathbf{Z}^*$ then $\sup_{z \in \mathbf{Z}} \Phi(c^*, z) = \infty$ while, if $c^* \in \mathbf{Z}^*$ then $\phi(c^*, z) = \mathcal{F}(c^*)$ by definition. We have, therefore, to show

$$\inf_{c^* \in \mathbf{C}^*} \sup_{z \in \mathbf{Z}} \Phi(c^*, z) = \sup_{z \in \mathbf{Z}} \inf_{c^* \in \mathbf{C}^*} \Phi(c^*, z) .$$

It is trivial that

$$\inf_{c^* \in \mathbf{C}^*} \sup_{z \in \mathbf{Z}} \Phi(c^*, z) \geq \sup_{z \in \mathbf{Z}} \inf_{c^* \in \mathbf{C}^*} \Phi(c^*, z) := \underline{I} ,$$

so we only have to show that

$$\inf_{c^* \in \mathbf{C}^*} \sup_{z \in \mathbf{Z}} \Phi(c^*, z) \leq \underline{I} . \tag{6.1}$$

Define, for any $z \in \mathbf{Z}$

$$A_z = \{c^* \in \mathbf{C}^* \ ; \ \Phi(c^*, z) \leq \underline{I}\} .$$

Note that (6.1) follows provided

$$\bigcap_{z \in \mathbf{Z}} A_z \neq \emptyset . \tag{6.2}$$

The next step is to show that, for any finite set $z_1, \ldots z_n \in \mathbf{Z}$, the set $\bigcap_{z_i} A_{z_i} \neq \emptyset$. The proof of this part can be taken from the proof of Theorem 2.8.1 in [Ba].

Finally, note that $A_0 \subset \overline{A}_0$ as defined in the Proposition, since $\underline{I} \leq I$. It follows that A_0 is compact, and that the non-empty intersection of finite sets implies (6.2). $\square$

Proof. of Lemma 4.7:
Let $\rho_1(r)$ be a smooth, positive function with compact support such that

$$|\mathbb{S}^{n-1}|\int_0^\infty r^{n-1}\rho_1(r)dr = 1 \ ; \quad \int_0^\infty r^k\rho_1(r)dr := M_k \ ; \quad \int_0^\infty r^{n-1}\rho_1^p dr := L(p) \ .$$

Set also

$$\rho_\alpha(r) = \alpha^n \rho_1(\alpha r) \ .$$

Define

$$\mathbf{v}(x,t) = \begin{cases} \frac{x-\bar{x}(t)}{t-t_0} + \dot{\bar{x}}(t) & \text{if } t_0 \le t \le (t_0+t_1)/2 \\ \frac{x-\bar{x}(t)}{t_1-t} + \dot{\bar{x}}(t) & \text{if } (t_0+t_1)/2 \le t \le t_1 \end{cases}$$

$$\rho(x,t) = \begin{cases} \frac{1}{(t-t_0)^n}\rho_\alpha\left(\frac{|x-\bar{x}(t)|}{t-t_0}\right) & \text{if } \ t_0 \le t \le (t_0+t_1)/2 \\ \frac{1}{(t_1-t)^n}\rho_\alpha\left(\frac{|x-\bar{x}(t)|}{t_1-t}\right) & \text{if } \ (t_0+t_1)/2 \le t \le t_1 \end{cases} .$$

A direct calculation shows that ρ satisfies the weak form of the continuity equation:

$$\rho_t + \nabla_x \cdot (\mathbf{v}\rho) = 0 \ .$$

Let us now consider the interval $[t_0, (t_0+t_1)/2]$. The second interval $[(t_0+t_1)/2, t_1]$ can be treated analogously. Define the lifting of ρ as

$$f(x,t,v) = \sigma^{-n}\pi^{-2/n}\exp\left(-\frac{|v-\mathbf{v}|^2}{\sigma^2}\right)\rho(x,t) \ .$$

It follows immediately that

$$\int_{\mathbb{R}^n} v^2 f(x,t,v)dv = \frac{\sigma n}{2}\rho(x,t) + \mathbf{v}^2\rho(x,t) \ ; \quad \int_{\mathbb{R}^n} |f|^p dv = p^{-n/2}\pi^{2(1-p)/n}\sigma^{n(1-p)}\rho^p(x,t) \ .$$

Moreover:

$$\int_\Omega \rho(x,t)dx = 1 \ ; \quad \int_\Omega \rho^p(x,t)dx = (t-t_0)^{n(1-p)}|\mathbb{S}^{n-1}|\int_0^\infty r^{n-1}\rho_\alpha^p(r)dr = \alpha^{n(p-1)}(t-t_0)^{n(1-p)}|\mathbb{S}^{n-1}|L(p)$$

$$\int_\Omega |\mathbf{v}(x,t)|^2\rho(x,t)dx = |\dot{\bar{x}}(t)|^2 + |\mathbb{S}^{n-1}|\int_0^\infty r^{n+1}\rho_\alpha(r)dr = |\dot{\bar{x}}(t)|^2 + |\mathbb{S}^{n-1}|\alpha^{-2}M_{n+1}$$

In particular:

$$\int_\Omega\int_{t_0}^{(t_0+t_1)/2} |v|^2 f = \int_{t_0}^{(t_0+t_1)/2} |\dot{\bar{x}}|^2\, dt + C_1|t_1-t_0|\alpha^{-2} + O(\sigma)$$

and

$$\int_\Omega\int_{t_0}^{(t_0+t_1)/2} \rho^p = C_3|t_1-t_0|^{n(1-p)+1}\alpha^{n(p-1)}$$

□

Proof. of Lemma 5.8
We use Corollary 4.1 with μ supported on $\Omega \times [t_1, t_0]$ and $\mu_{t_0} = \delta_{x_0}$, $\mu_{t_1} = \delta_{x_1}$, to obtain

$$\phi(x_1, t_1) - \phi(x_0, t_0) \leq \frac{1}{2}||\mu||_2^2 + \left| \int_\Omega \int_{t_0}^{t_1} (\phi_t + |\nabla_x \phi|^2/2)\mu_{(t}(dx)dt \right| . \tag{6.3}$$

By Lemma 4.7 we can find such a μ for which:

$$||\mu||_2^2 \leq \int_{t_0}^{t_1} |\dot{\overline{x}}|^2 + C_1|t_1 - t_0|\alpha^{-2} \tag{6.4}$$

and, for the density $\rho = \rho_\mu$:

$$\int_\Omega \int_{t_0}^{t_1} \rho^p \leq C_3|t_1 - t_0|^{n(1-p)+1}\alpha^{n(p-1)} \tag{6.5}$$

where $p < 1 + 1/n$ and α any positive constant. Since ρ is supported, for any t, in a domain of diameter $(t_1 - t_0)\alpha^{-1}$ it follows

$$\int_\Omega \int_{t_0}^{t_1} (\phi_t + |\nabla_x \phi|^2/2)\rho dx dt = O\left(\frac{L(t_1 - t_0)^2}{\alpha}\right) + \int_{t_0}^{t^1} P(\overline{x}(t), t)dt + \int_\Omega \int_{t_0}^{t_1} \xi \rho dx dt ,$$

where L is the Lipschitz norm of P. By (6.5) we obtain

$$\left| \int_\Omega \int_{t_0}^{t_1} \xi \rho dx dt \right| \leq ||\rho||_p ||\xi||_s \leq C_3^{1/p}|t_1 - t_0|^{[n(1-p)+1]/p}\alpha^{n(p-1)/p}||\xi||_s . \tag{6.6}$$

Collecting (6.3) to (6.6)

$$\phi(x_1, t_1) - \phi(x_0, t_0) \leq \frac{1}{2}\int_{t_0}^{t_1} |\dot{\overline{x}}|^2 dt + C||\xi||_s|t_1 - t_0|^{[n(1-p)+1]/p}\alpha^{n(p-1)/p} + C_1|t_1 - t_0|\alpha^{-2} + O\left(\frac{L(t_1 - t_0)^2}{\alpha}\right) .$$

The choice $\alpha = ||\xi||_s^{-\beta}(t_1 - t_0)^\gamma$ where $\gamma = \frac{(n+1)(p-1)}{2p+n(p-1)}$ and $\beta = \frac{p}{2p+n(p-1)}$ is the optimal choice and yields the desired result. $\square$

References

[Am] L. Ambrosio: *Lectures Notes on Optimal Transport Problems*, CVGMT Preprint: http://cvgmt.sns.it/papers/amb00a/

[AGS] L. Ambrosio, N. Gigli & G. Savare: *Gradient flows of probability measures*, Preprint.

[AFP] L. Ambrosio, N.Fusco & D.Pallara: *Functions of Bounded Variations and Free Discontinuity Problems*, Oxford University Press, 2000.

[B] Y. Brenier: *Polar factorization and monotone rearrangement of vector valued functions*, Comm. Pure Appl. Math, **44**, (1991), 375-417.

[Ba] A.V. Balakrishnan, *Applied Functional Analysis*, Applications of Mathematics **3**, Springer-Verlag, 1976.

[BB] J.D.Benamou, Y. Brenier: *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer.Math., **84** (2000), 375-393.

[BBG] J.D.Benamou, Y. Brenier & K.Guitter: *The Monge-Kantorovich mass transfer and its computational fluid mechanics formulation*, Inter. J. Numer.Meth.Fluids, **40** (2002), 21-30. (ed. by Talenti), (1996), Dekker

[E] L.C.Evans, *Partial Differential Equations* 1949- Providence, R.I. : American Mathematical Society, c1998.

[M] G. Monge: *Mémoire sur la théorie des déblais et de remblais*, Histoire de l'Académie Royale des Sciences de Paris, 1781, pp. 666-704

[RKF] J. Richard, D. Kinderlehrer & F. Otto: *The variational formulation of the Fokker-Planck equation.* SIAM J. Math. Anal. **29** (1998), no. 1, 1–17

[V] C. Villani: *Topics in Optimal Transportation*, Graduate studies in Math, **58**, AMS, 2003

[W] G. Wolansky: *Rotation numbers for measure valued circle maps*, J. D'Anal. Math., **97**, 169-201, 2005

[W1] G. Wolansky: http://arxiv.org/abs/math-ph/0306070